COMMENTS

On Testing the Connection Between Economic Freedom and Growth,
Robert A. Lawson 398-406

How to Handle Economic Freedom: Reply to Lawson,
Jakob de Haan and Jan-Egbert Sturm 407-411

Assume the Positional: Comment on Robert Frank,
Andrew Kashdan and Daniel Klein 412-434

Taking Libertarian Concerns Seriously: Reply to Kashdan and Klein,
Robert H. Frank 435-451

Work Incentives and Employment are the Wrong Explanation of Sweden's Success,
Andreas Bergh 452-460

Second Reply to Bergh, Peter H. Lindert 461-465

A Little More Liberty: What the JEL Omits in Its Account of what
The Economic Report of the President Omits,
Daniel Klein and Michael J. Clark 466-483
Benjamin Franklin and Colonial Money: A Reply to Michener and Wright—Yet Again, *Farley Grubb* 484-510

**CHARACTER ISSUES**

Knut Wicksell, Gustav Cassel, Eli Heckscher, Bertil Ohlin, and Gunnar Myrdal on the Role of the Economist in Public Debate,
*Benny Carlson and Lars Jonung* 511-550

**DO ECONOMISTS REACH A CONCLUSION?**

Do Economists Reach a Conclusion on Rail Transit?
*Ted Balaker and Cecilia Joung Kim* 551-602

**CORRESPONDENCE** 603-604
On Testing the Connection between Economic Freedom and Growth

ROBERT A. LAWSON*


Abstract

JAKOB DE HAAN, SUSANNA LUNDSTRÖM, AND JAN-EGBERT Sturm have written a valuable survey of the literature that uses the Gwartney and Lawson economic freedom (EFW) index. Their discussion of the index’s theoretical underpinnings and methodological ins and outs itself should be useful to scholars interested in the field.

Also their survey of the empirical economic freedom–economic growth literature is a useful contribution.

It is clear from these studies that EF [economic freedom] seems to have a positive association with growth. None of the studies summarized reports that economic freedom is bad for growth. Even though many studies have serious drawbacks . . . it is a strong result that emerges when

* Department of Economics, Capital University. Columbus, Ohio 43209-2394.
This paper was supported by the Social Philosophy and Policy Center at Bowling Green State University.
looking at these studies collectively. Furthermore, those studies that deal with the problems of model specification and sensitivity in a more rigorous way also find that there is a positive growth effect from EF. (170)

While De Haan et al. accurately describe the mechanics of the construction of the EFW index and the econometric literature that has found a link between EF and economic growth, I find myself in disagreement with some of their commentary. This comment, in part, will address these issues.

ECONOMIC FREEDOM AND IDEOLOGICAL BIAS

De Haan et al. (158-159) echo Paldam’s (2003) claim that the authors of the EFW index are “zealots”. But their remarks are not unfriendly, and seem to recognize the virtues of scientific purpose combined with open ideological disclosure. The continued use of the term zealot, however, suggests that the EFW index authors will do or say anything to advance their cause. This is certainly not the case. While it is true that the authors and the publishers are ideologists in the sense that they are motivated by a set of ideas, this is true of all scientists whether they choose to admit it or not.

The desire to define and measure economic freedom is like the desire in the early twentieth century to define and measure GDP. At the time, many scholars were skeptical of the attempt to measure GDP and there were heated debates about methodological questions (like for example the treatment of government spending in GDP). Were the measurers of GDP driven by an ideology? Yes. Were they zealots? I think not. They were scientists who sat down to design, as best they could with the tools at hand, a measure of the current economic activity of the nation. Economic activity exists and their job was to measure it.

Likewise economic freedom exists. It is a thing. We can define and measure it. Although the research literature has shown many desirable consequences associated with economic freedom, there is nothing in the

1 Klein (2006) explores these virtues in connection with the advice of Gunnar Myrdal.
EFW index project that presumes freedom is a good thing or that we should necessarily have more of it.

A primary purpose for the creation of the EFW index was to inject some much needed scientific fact into the ongoing debate about the merits of free-market economic systems versus interventionist systems. What had characterized this debate for most of its history was a paucity of data and evidence. With the creation of the EFW index we are now in a position to begin to address the problem of economic organization as scientists should by measurement of reality and testing of various hypotheses.

**ECONOMIC FREEDOM AND THE ROLE OF GOVERNMENT**

De Haan et al. (163-164) mildly criticize the EFW index’s treatment of government taxation both in the traditional sense and in the case of the “inflation tax.” The EFW index separately treats government spending/taxation and inflation, each inversely affecting the index ratings. That is, larger government spending, higher taxes, and higher inflation result in lower index ratings.

De Haan et al. note, as do the EFW index authors, that there are many arguments for government spending/taxation and inflation from various perspectives: economic growth, macroeconomic stability, efficiency, or income redistribution, to name just a few. But the EFW index is an economic freedom index—it is not an index of economic growth policies, efficient government provision of public goods, macroeconomic stabilization policies, or ideal income distribution policies.

The purpose of the EFW index is to measure, no doubt imprecisely, the degree of economic freedom that exists. From the perspective of individual economic freedom, government taxation (whether the traditional form or via inflation) is a substitution of the “collective will” for the individual’s, is an expropriation of private resources, and is a violation of individual economic liberty from the very first dollar.

Of course, it may be the case that there is an optimal level of economic freedom, especially when economic freedom conflicts with other goals. Trade-offs between liberty and security or other values may exist, and some countries may have too much economic freedom, like too much air pollution. The EFW index itself is agnostic about whether there is an optimal level of economic freedom or what that level might be. The EFW index simply measures.
Furthermore, if there is a positive role to be played by government with respect to economic freedom, perhaps through the enforcement of private contracts or private property enforcement that requires taxation, then other areas of the EFW index will capture those positive impacts. For example, a country with no government taxation or spending would score very high in the government size area, but could score very low in the property rights area (assuming no private property rights protections were put in place).

DECOMPOSITION OF THE PARTS OF ECONOMIC FREEDOM: WHAT MATTERS MOST?

In recent years a literature has grown up that attempts to address the question of which of the various parts of the EFW index matter the most for economic growth. Although this seems like a worthwhile project, this is like asking what the most important ingredient is in a cherry pie. Is it the cherries? The sugar? The flour? The shortening? If you fail to include any of those ingredients, you will not bake a cherry pie. In reality all the ingredients work together. While there is some room for substitutions and variations (e.g., butter instead of shortening), you have to be careful, as not all substitutions will work (e.g., salt for sugar). You could even substitute apples for cherries, in which case you can still get a tasty pie, even though no longer a cherry pie. The point is that the relationship between EF and economic growth is complex. We can parse EF out into various parts, ingredients if you will, but it is conceptually difficult to say which is most important.

Econometrically, this issue presents itself as a multicollinearity problem. Many of the component parts of the EFW index are highly correlated with other parts. Simply breaking up the index into parts and running growth regressions results in such massive multicollinearity that the coefficient estimates on individual components become meaningless. Dropping out components to reduce this introduces the problem of omitted variables.

The bottom line is that the current measures of EF are too crude, and our econometric methods too imperfect, to answer these questions satisfactorily. Most of the “what matters most” literature has failed to
acknowledge these issues. In stark contrast to the rest of the paper, De Haan et al. (168-170) survey this literature without much criticism.

**WHAT MATTERS FOR ECONOMIC GROWTH: ECONOMIC FREEDOM LEVELS OR CHANGES?**

De Haan et al. (170, 176-177) criticize a number of studies for including in growth regressions both the level of EF and the change in EF. They advocate including only the change, going so far as to say that including the level is “wrong” as a matter of statistical methodology. I disagree.

Consider the following specifications:

(1) \[ \text{GROWTH} = \alpha_1(\text{EF}_1) + \alpha_2(Z) \]

(2) \[ \text{GROWTH} = \beta_0(\text{EF}_0) + \beta_1(\text{EF}_1) + \beta_2(Z) \]

Where GROWTH is the rate of economic growth over some period of time. EF\(_0\) is the economic freedom level at the beginning of the period and EF\(_1\) is the level of economic freedom at the end of the period. Z is a matrix of control variables (including the constant).\(^2\)

Clearly, Equation (1) is inappropriate as it is logically impossible for the level of economic freedom at the end of the period to affect economic growth in the previous period. Something occurring today cannot determine what happened yesterday.

Equation (2) on the other hand is different. If you include both the level of EF at the beginning of the period and at the end of the period, the coefficient \(\beta_1\) measures the impact of the level of economic freedom at the end of the period after controlling for the level at the beginning of the period. If you hold the initial level of EF constant by including it as a variable, the end level of EF is a result of the change during the period. Thus, in Equation (2)

\(^2\) At a minimum control variables would normally include the level of income at the beginning of the period and measures of the rate of investment in physical and human capital.
the coefficient $\beta_1$, in effect, measures the change in EF since the beginning of the period. And we all agree that the change in EF over the period may impact growth.

I agree that Equation (2) is still a problematic specification because the level of EF at the beginning of the period is likely to be highly collinear with the level of EF at the end of the period. This high degree of collinearity between EF$_0$ and EF$_1$ will make the coefficient estimates in Equation (2) difficult to interpret and will bias the standard errors.

An equivalent specification would be to include the EF level at the beginning of the period and the change over the period, as in Equation (2').

$$(2') \text{GROWTH} = \gamma_0(\text{EF}_0) + \gamma_1(\Delta \text{EF}) + \gamma_z(Z)$$

Where $\Delta \text{EF} = \text{EF}_1 - \text{EF}_0$. Equations (2) and (2') are functionally equivalent as there is no more statistical information in Equation (2') than in Equation (2). Equations (2) and (2') will have identical $R^2$'s and the $\beta_i$ and $\gamma_i$ coefficients will be algebraic functions of each other: Specifically, $\beta_0 = \gamma_0 - \gamma_1$, and $\beta_1 = \gamma_1$. But since there is less of a collinearity problem in Equation (2') between EF$_0$ and $\Delta$EF, the coefficient estimates are easier to interpret and the test statistics unbiased.

De Haan et al. argue that these specifications are inappropriate and invite “causality and endogeneity issues” (176) because Equation (2) and thus implicitly Equation (2') include the level of EF at the end of the period. I argue here that including the level of EF at the end of the period is not a problem so long as you control for the level of EF at the beginning of the period, thus in effect converting the end of the period variable into a change in EF measure.

In contrast, De Haan et al. advocate a specification in which only the change in EF is included as in Equation (3).

$$(3) \text{GROWTH} = \zeta_1(\Delta \text{EF}) + \zeta_z(Z)$$

---

$^3$ Similarly, Equation (2) is equivalent to a specification that includes the change in economic freedom and the level of economic freedom at the end of the period as in Equation (2'').

$$(2'') \text{GROWTH} = \theta_0(\text{EF}_1) + \theta_1(\text{EF}_1 - \text{EF}_0) + \theta_z(Z)$$
It should be noted that their preferred specification also includes \( \Delta EF \) at the end of the period. Recall that \( \Delta EF = EF_1 - EF_0 \) thus Equation (3) can be rewritten as Equation (3').

\[
(3') \text{GROWTH} = -\zeta_1(\text{EF}_0) + \zeta_1(\text{EF}_1) + \zeta(Z)
\]

Equation (3') is nothing more than a restricted version of Equation (2) where \( \beta_1 = -\beta_0 \) in Equation (2)—or equivalently that \( \gamma_0 = 0 \) in Equation (2'). It is important to note that the validity of this restriction is a testable hypothesis. The standard t-statistic for \( \gamma_0 \) in Equation (2') would tell us whether that coefficient is statistically different from zero. The estimates presented in Table 6 of their article clearly show that \( \gamma_0 \) is statistically significant.\(^4\) That is, the hypothesis that we should drop the level and include only the change in economic freedom can be rejected statistically.

Alternatively (and perfectly equivalently) we can compare the Sum of Squared Residuals (SSR) of Equations (2) and (3) to test this hypothesis. De Haan et al.’s estimate for the SSR in Equation (2) is 361.38 and the SSR in Equation (3) is 396.64. The calculated F-statistic for the null hypothesis \( H_0: \beta_1 = -\beta_0 \) is \( F_{1,74} = 6.97 \) (\( p = 0.01 \)). Thus we statistically reject the null hypothesis and conclude that \( \beta_1 \neq -\beta_0 \); that is, again, their restriction cannot be justified on statistical grounds.\(^5\)

Furthermore, I would argue that the empirical specification that includes both the EF level and change is more defensible. Hong Kong has been at the top of the EFW index for all years since 1970. But it has shown little or no change in its rating—it is at the top and can not go up. In contrast, countries like Nicaragua have shown large improvements in their index ratings in recent years but still have relatively low ratings. Would we really expect Nicaragua to grow faster than Hong Kong simply because its rating has increased a lot while Hong Kong’s has not? By omitting the level of economic freedom from the specification, De Haan et al. fail to acknowledge the possibility that a country with a high degree of economic freedom is likely to outgrow a country with low economic freedom. Countries like Hong Kong would end up being outliers in De Haan et al.’s approach.

Both the EF level and change appear to impact economic growth. There is no statistical reason to omit level in favor of just change.

\(^4\) See Regression (2) in De Haan et al.’s Table 6.
\(^5\) See Regression (3) in De Haan et al.’s Table 6. Thanks to Jan-Egbert Sturm for providing the underlying regression statistics from Table 6 of their paper.
A CRITICISM OF SPECIFICATION TESTING

De Haan et al. (177) are highly critical of much of the EF-growth literature for failing to adequately check on the sensitivity of their results. I am in complete agreement with De Haan et al. that specification matters and that it matters a lot. Using both economic and statistical theory as our guide, we need to think carefully about how we specify empirical models. Data mining, the practice of almost randomly including and excluding variables in an empirical model, to get desired results is a serious problem.

De Haan et al. advocate one of the various specification-testing approaches, one in which a large number of possible regressions are run to examine the range of parameter estimates that can result. My criticism is this: what has happened to theory? For example, standard economic theory says that investment is an important factor for growth. We know then that failing to include investment in a regression will result in biased estimates of the remaining parameters. Why should we pay any attention to such an obviously mis-specified model? De Haan et al. sharply criticize others for failing to run the “right” specification, i.e., one that includes changes in economic freedom only but omits the level of economic freedom, and then go on to say researchers should intentionally run hundreds and maybe thousands of “wrong” specifications.

Certainly, we should recognize that ambiguity in economic and statistical theory requires us to run different reasonable versions of our models to explore the sensitivity and robustness of our parameter estimates. I am suspicious of any empirical research that presents the one right model. But I am also suspicious of someone who runs hundreds of meaningless models and then claims the “results” have any meaning. In my view, turning the research process over to a machine is worse than the data mining crime that it is meant to combat.

CONCLUSION

De Haan et al.’s review article is an important contribution to the literature on economic freedom and economic growth, and will become the starting point for future researchers in the field. My comments are intended
as inviting possible clarifications to an article that remains very valuable regardless of possible shortcomings.

REFERENCES


ABOUT THE AUTHOR

Robert A. Lawson is Professor of Economics and George H. Moor Chair in the School of Management at Capital University in Columbus, Ohio. He earned his B.S. in economics from the Honors Tutorial College at Ohio University and his M.S. and Ph.D. in economics from Florida State University. Lawson is co-author of the *Economic Freedom of the World* reports and has professional publications in *Public Choice, Cato Journal, Kyklos, Journal of Labor Research, Journal of Institutional and Theoretical Economics*, and *European Journal of Political Economy*. He is a senior fellow with the Buckeye Institute for Public Policy Solutions where he has written extensively on issues of state and local public finance. Lawson served as president of the Association of Private Enterprise Education and is a member of the Mont Pelerin Society. Email: rlawson@capital.edu.

[GO TO REPLY BY DE HANN AND STURM](#)

ECON JOURNAL WATCH 406
How to Handle Economic Freedom: Reply to Lawson

JAKOB DE HAAN* AND JAN-EGBERT STURM**

In his comment on the Economic Freedom Survey article we wrote with Susanna Lundström (De Haan et al. 2006), Lawson (2006) criticizes five elements of our article:

1. Our use of the term “zealots”
2. Our remarks on the role of government in the construction of the economic freedom index
3. Our discussion of the composition of the economic freedom index
4. Our criticism of empirical growth models that include both the level and the change in economic freedom
5. Our endorsement of a modeling approach wherein researchers run many specifications

In this reply, we briefly deal with each of these issues.

As Lawson points out, we cite Martin Paldam (2003) to convince skeptical readers that the economic freedom indicators provide valuable information and should be used more often in empirical research. Indeed, the change in the economic freedom indicator can be used as a proxy for the institutional changes and policies generally supported by the IMF. The economic freedom index, therefore, helps us to examine whether the IMF is right when it imposes conditions on loans to member states. This and other highly relevant issues can be fruitfully approached using the economic

* University of Groningen. Groningen, The Netherlands 9700AV.
** KOF—Swiss Institute for Business Cycle Research, ETH. Zurich, Switzerland 8092.
freedom index. So we endorse Lawson’s view that “With the creation of the EFW index we now are in a position to begin to address the problem of economic organization as scientists should by measurement of reality and testing of various hypotheses” (Lawson 2006, 400).

Why then, has the index not been used more often? We suspect that this may be due to the pro-free market position of institutes like the Fraser Institute and the Heritage Foundation, which publish the indexes. We argue that this is not a good reason to neglect the indexes, for reasons expressed in the passages we quoted from Paldam. We fully agree with Paldam when he refers to Lawson and others as a group of eminent scholars.

As to our criticism of the way government taxes are taken up in the economic freedom index, let us repeat what we wrote in our survey: “Of course, taxes always distort prices, but that in itself does not make it necessary to include the level of taxation in an index of economic freedom. To extreme libertarians, like Rothbard (1970), the state obtains its revenue by coercion, known as taxation, whereas private persons and groups obtain their income voluntarily by selling goods and services to others or by voluntary gift. To Rothbard, taxation is theft, pure and simple” (De Haan et al. 2006, 164). Lawson affirms that taxation is coercion and hence an incursion on economic freedom. In our view, that interpretation goes too far, as economic freedom also implies that government has an important role to play, notably in securing property rights. However, Lawson is right, of course, that the index does elsewhere capture these freedom-enhancing effects of government.

On the composition of economic freedom, we have a more substantial disagreement with Lawson. The development of the economic freedom index shows that even when one agrees on the general definition of economic freedom, there are still many options when it comes to constructing an index. In Lawson’s analogy, there is not a unique recipe for a cherry pie. However, there is an interesting discussion of the proper ingredients. Various authors have pointed out that many of the candidate ingredients, which are supposed to reflect a particular dimension of economic freedom, have a low correlation with other ingredients and with the aggregate index. This may suggest that the various elements of economic freedom may not be measuring the same thing. In other words, there are many proxies for the latent variable economic freedom. In our view, this calls for using latent variable techniques both when it comes to clustering particular elements of economic freedom as well as aggregating these elements into (one or more) aggregates. In our view, the current clustering of the various elements of the economic freedom into five categories, like size of government and sound money, is rather arbitrary. In our view, it makes more sense to use, for
instance, factor analysis to decide on the number and composition of the various categories. These can then, in turn, be employed in empirical growth models to check to what extent different dimensions of economic freedom may differently affect income growth. Our favored approach implies that we share Lawson’s skepticism of studies that examine the relationship of the individual elements of economic freedom and economic growth. We thank Lawson for offering us the opportunity to point this out.

We also disagree with Lawson on the issue of using both the level and the change in economic freedom. We all agree that Lawson’s equation (1) does not make theoretical sense: the level of economic freedom at the end of the sample period cannot explain economic growth experienced over the sample. It would rather indicate that causality runs the other way. We also all agree that the change in economic freedom may impact economic growth. However, such possible impact is something that should be tested for, using appropriate econometric modeling approaches that take the problem of model uncertainty into account, such as the Extreme Bounds Analysis that we apply in much of our research. We also agree that equations (2) and (2’) are mathematically equivalent. This, however, also implies that both specifications suffer from the same kind of multicollinearity problems and hence that it does not matter econometrically which specification is estimated (the estimation results clearly show this). If, however, the estimate for β₁ is larger than that for β₀ (i.e. γ₀<2γ₁) the danger of having estimated an equation with reverse causality problems (as in equation (1)) is substantial and not to be neglected. We hence disagree with Lawson’s statement that “including the level of EF at the end of the period is not a problem so long as you control for the level of EF at the beginning of the period, thus in effect converting the end of the period variable into a change in EF measure” (403). It depends upon the estimation results. Column (3) of Table 6 in our paper actually shows a negative estimate for β₀ and a high and significant estimate for β₁ (or, in Column (2) of the same table, an estimate of γ₀ which is significantly less than twice the estimate of γ₁). As we are not interested in the effect of growth on institutional reform in our paper, we (at least to a certain extent) circumvent this reverse causality problem by allowing into the specification only either economic freedom at the beginning of the period or the change in economic freedom. When we include only the level of economic freedom in the regression it has no significant impact (see column (4) in our Table 6). The only reason that it is significant in column (2) of Table 6 is that in that column the change in economic freedom is included. Hence, equation (1) of Lawson and his preferred specification (2’) that includes both the change and the initial level of
economic freedom are statistically equivalent. If equation (1) is to be considered nonsensical, then—given the estimation results—Lawson’s proposed equation is as well, and for the same reason.

Finally, we do not agree with Lawson’s criticisms concerning our favored modeling strategy, the Extreme Bounds Analysis (EBA). Lawson (405) states, “My criticism is this: what has happened to theory? For example, standard economic theory says that investment is an important factor for growth. We know then that failing to include investment in a regression will result in biased estimates of the remaining parameters. Why should we pay any attention to such an obviously misspecified model?” If Lawson feels, like we do, that a good case can be made to include investment, then there is no reason not to do so. In fact, in our papers on the relationship between economic freedom and economic growth we always included investment in the so-called M-vector of the EBA, i.e. it is included in all regressions (see, for instance, De Haan and Sturm 2000). Likewise, we always include initial income and often also a human capital variable. At the same time, the inclusion of these variables has been criticized by other authors. Fortunately, the EBA is flexible enough to deal with this: if the researcher feels—on theoretical or empirical grounds—that a particular variable should not be included in the M-vector she can easily move it to the Z-vector.

CONCLUSION

Even though our reply might suggest otherwise, there are many issues on which Lawson and we agree, like the usefulness of the economic freedom index, the need for solid research on the relationship between this index and economic growth, and the importance of academic debate. We believe this exchange is a good example of academic debate as a discovery procedure.

REFERENCES


ABOUT THE AUTHORS

Jakob de Haan is Professor of Political Economy, University of Groningen, the Netherlands. He is also director of SOM, the research institute and graduate school of the faculties of Management and Organisation and Economics of the University of Groningen. He is editor of the European Journal of Political Economy (Link), and President of the European Public Choice Society (Link). His email address is Jakob.de.Haan@rug.nl.

Jan-Egbert Sturm is Professor of Applied Macroeconomics and director of KOF—Swiss Institute for Business Cycle Research both at the ETH Zurich, Switzerland (Link). He is also a member of the European Economic Advisory Group at CESifo, Munich (Link). His email address is sturm@kof.ethz.ch.

GO TO COMMENT BY LAWSON (2006)
Assume the Positional: Comment on Robert Frank

ANDREW KASHDAN AND DANIEL B. KLEIN

Abstract

The individual spectator in a stadium can improve his or her view by standing on tiptoe. But when everyone does likewise, everyone is worse off. As Fred Hirsch (1976, 5) notes when introducing this example, “the preferred outcome may be attainable only through collective action,” which could mean an explicit or implicit agreement. Economists will recognize it as a prisoner’s dilemma. Hirsch and Robert Frank extend the idea to a wide variety of consumption decisions involving what Hirsch called positional goods, whose value depends relatively strongly on position in relation to others. Certain luxury goods are the classic examples, but other prominent examples include education and certain jobs or leadership positions. Essentially, Frank sees competition for positional goods as wasteful, and wants higher taxes to discourage it. We argue that he overstates the problem, overlooks various voluntary solutions, overlooks unintended consequences of using taxation, and neglects the Smithian incumbency on those proposing coercion.

* Department of Economics, George Mason University, Fairfax, VA 22030. The authors would like to thank workshop participants at the Mercatus Center at George Mason University.
We have great admiration for Robert Frank—for his breadth of insight, his industrious service to economic education, his focus on important policy issues, his exemplary participation in public discourse, his habits of plain language, and his important illumination of issues of status and positionality (and here his place is pre-eminent!). We regret, however, that he does not combine these virtues with more libertarian sensibilities about politics, government, and society.¹

Frank summarized his work on positional goods in an *American Economic Review* article “Positional Externalities Cause Large and Preventable Welfare Losses” (2005). As in his books *Choosing the Right Pond* (1985a), *The Winner-Take-All Society* (Frank and Cook 1995) and *Luxury Fever* (1999), Frank argues that positional goods are creating negative externalities and inefficiency that can be reduced by higher taxes. In an earlier paper on positional externalities, Frank argued that “The apparent strength of consumption externalities suggests that supply-siders are barking up the wrong tree when they say that income and consumption taxes introduce distortions into important economic decisions. Rather, such taxes alleviate existing distortions in those same decisions” (Frank 1985b, 102).

Frank follows Hirsch in his conception of the problem, and Hirsch has a clear discussion of the issues at stake. A public-goods element can be attributed to a wide range of private expenditures. Hirsch writes the following of education and job seeking: “the utility of expenditure on a given level of education as a means of access to the most sought after jobs will decline as more people attain that level of education” (Hirsch 1976, 3). Since the benefits of education depend at least partly on relative position rather than absolute skills or learning, “the possibility of a general advance is an illusion,” and “the distributional struggle returns” (Hirsch 1976, 6). Both Hirsch and Frank emphasize that at least some positional externalities are independent of psychological considerations or envy.² What people get depends to an increasing extent on their relative position in the economic

---

¹ Incidentally, Frank (2003, 2) claims that his tax plan accords with the “classical libertarian criterion” that the state may interfere in actions that bring “harm to others.” But Frank should know that the liberty idea is rooted in property ownership (and freedom of contract), not guarantees against any kind of harm. Just as a competitor might harm you, your neighbor’s ownership of a fancy car might harm you because of positional rivalry. But under no jurisprudence, much less a classical liberal jurisprudence, does it tread on any of your property. Moreover, as a supposed punishment or deterrent of positional crimes, Frank’s taxation is indiscriminate, as it brings force against the guilty and innocent alike.

² See Frank (1985a, 114; 2003, 13-14; 2005, 141) and Hirsch (1976, 8).
hierarchy, a factor separate from the level of satisfaction in their consumption.

Another way to think of the concept of positionality is that certain goods are characterized by some type of absolute scarcity, because of physical or social limits (hence Hirsch’s "Social Limits to Growth"). A particular natural landscape—the heights—is physically scarce, and leadership positions are socially scarce. All economic goods are scarce, but, while the supply of regular goods can be increased by production over time, supplies of other goods are fixed (Hirsch 1976, 22). Builders can build more skyscrapers, but there can only be one tallest building. In such cases, free competition will involve wasteful social climbing. For highly positioned jobs, the credentials required might increase, and the resulting resource utilization might be wasteful overall.

Frank advocates taxation to reduce spending on positional goods and encourage consumption of less positional goods, such as vacation time (2005, 137). The exact form of taxation has varied in Frank’s writings. He advocated taxing specific categories of consumption goods in Choosing the Right Pond (1985a, 231-232), a general progressive consumption tax in The Winner-Take-All Society (1995) and Luxury Fever (1999), and a steeper progression in the income tax in an unpublished paper (2003). We suspect that Frank (2003) returns to an income tax, even though he earlier (1999) gave reasons for the superiority of consumption taxes, because it is more feasible politically. He continues to note the objective of “raising the price of consumption relative to leisure” (2003, 25).

According to both Hirsch and Frank, the reason that such externalities have become increasingly important is that the social “heights” remain fixed while the material purse continues to grow, increasing relative demand for positional goods. Hirsch published his book in 1976 and wrote that the positional economy had not become a significant issue until “recent times.” Frank’s Luxury Fever was published in 1999, and mostly focuses on spending patterns of the 1990s. Much earlier, Thorstein Veblen wrote about “conspicuous consumption” partly in response to the houses built by millionaires in the late 19th century, and in the same period Walter Weyl wrote of “new types of destitutes—the automobileless, the yachtless, the Newport-cottageless.” Still earlier, Adam Smith wrote about the poor man

---

3 As in Frank (1985a), Hirsch favors more targeted policy solutions, including reducing pay for positional jobs and having government allocate positional goods (Hirsch 1976, 183-185).

4 Quoted in Hofstadter (1955, 147). Frank claims the Gilded Age was different because most families still had to tend to their basic material needs and the latest boom affects a much larger number of people (1999, 15).
who sees the conveniences of the rich, and who, “devotes himself for ever to the pursuit of wealth and greatness . . . at last to find [they are] mere trinkets of frivolous utility” (Smith 1982, 260-61). It is likely that, in one form or another, satiation of non-positional material goods and the onset of positional strivings have been claimed for other periods as well.

For Frank, increased income inequality means that the rich are spending more money on luxury and other positional goods, and this spending affects the group slightly below them, setting off an “expenditure cascade.” In order to keep up, the middle class must spend more and, therefore, work more hours and save less. Such a confluence of factors describes the “luxury fever” of the 1990s. “The basic question,” Frank writes, “is whether society should attempt to limit the consumption expenditures of top earners in the interest of constraining expenditure cascades that impose costs on families farther down the income scale” (2003, 23).

To head off objections about the illegitimacy of envy, Frank points out that “Expenditure cascades would continue to occur and would continue to impose large and tangible economic costs on middle-income families even if everyone were completely free of envy” (2003, 28). One of Frank’s prominent examples is spending on a bigger house to give your child a place in a better school district, which then leads to his or her gaining access to the top jobs. Because positions in a top school are scarce, the externality would occur even if parents were not concerned with comparing the house itself to those of their neighbors. The result is a prisoners’ dilemma:

---

5 Income inequality in turn, according to Frank, has been exacerbated by the growing importance of winner-take-all markets. In these markets, small differences in relative performance give rise to large differences in income, increasing the degree of scarcity for top positions (Frank and Cook 1995). We will not address this aspect of Frank’s argument in detail, but we can see that, to the extent he and Cook overstate the existence of winner-take-all markets, competition for these scarce jobs will be less wasteful. Cowen (2000, ch. 5) has incisively criticized Frank and Cook on winner-take-all markets in the context of top performers in media, sports and other areas of achievement. Cowen argues that in such contexts the winner-take-all aspect is diminishing rather than increasing. Cowen focuses on fame seeking, but many of his criticisms have broader application to Frank’s fretting about positional “arms races.”

6 Incidentally, Benjamin Friedman (2005, 93) argues that peoples’ aspirations for their children are likely to have an even more powerful effect on their behavior than aspirations for themselves.
When each family saves less to buy a house in a better school district, the net effect is merely to bid up the prices of those houses. Students end up at the same schools they would have attended if all families had spent less. (Frank 1999, 159)

And waste results in choosing to work more hours:

The person who stays at the office two hours longer each day to be able to afford a house in a better school district probably has no conscious intention to make it more difficult for others to achieve the same goal. Yet that is an inescapable consequence of his action. The best response available to others may be to work longer hours as well, thereby to preserve their current positions. (Frank 1999, 270)\(^7\)

But even if Frank is right that some positional striving does not involve envy, much of it does and is perceived to do so, so Frank’s imposition would tend to liberate and embolden envy as a political force.

**EXPLAINING MARKET FAILURE**

Can excessive spending to get your child into a top school be considered a market failure? Part of what makes education valued on a relative basis is that schools provide a signal apart from the actual skills and knowledge learned (Frank 1985a, 195-96; Spence 2002). Yet government interventions including compulsory attendance, free government schooling, and tough restrictions on opening and operating a private school have worked to lock in the reliance on the standardized signal and crowd out and hobble private solutions. Education delivers benefits beyond that of signaling, benefits that the substitutes to the top government schools can deliver. Home-schooled students compete very successfully for Ivy League

\(^7\) For other mentions of the housing and school district example, see Frank (2003, 13; 1995, 142).
spots. Perhaps the benefits of going to a highly ranked government high school derive largely from government intervention, and freedom would lead to less signaling waste and more real gains from education.

Frank acknowledges distortions created by government subsidies when it comes to state universities competing for elite status (Frank and Cook 1995, 222), but doesn’t extend the idea to the broader education market. Frank descends into pulp economics when addressing voucher proposals, saying that they might unleash “an educational rat race of unprecedented proportions” (1985a, 196). He apparently means that parents would have more purchasing power to spend on private school tuition. Frank fails to consider the competitive aspect of a private school system in comparison to a standardized government system. Private schools disciplined by the voluntary funding decisions of customers are likely to deliver a range of options, just as the home-schooling success illustrates in the current system. Frank’s argument for preventable externalities is about resources being wasted, not about the fairness of the distribution of income or consumption. Yet Frank neglects the inefficiency of a uniform system in which taxpayers pay for schools that their children do not well fit into. If schooling were debundled from residence and politics, the major example of positional “arms race” would largely disintegrate. But Frank never looks to depoliticization for a solution.

Neither is Frank’s argument for inefficient luxury spending convincing. If the standard for luxury spending has risen, so that the rich must buy ever more expensive cars to achieve elite status, the amount of waste may not be very significant. If status is the motivating factor, then much of the premium they pay ends up as producer surplus. Similarly, in the housing example, it is not clear how much of the “arms race” is actually wasteful. (Frank routinely writes of expenditure or positional “arms races,” as though social standing were a matter of war.) The bidding up of housing prices is not wasteful in itself, since those resources are not lost. Frank never dwells on how the price increase ought to induce new houses and new schools, nor does he acknowledge the waste arising from government land-use and housing restrictions. The waste, according to Frank, is the misallocation of the consumer’s labor hours, and the various decisions she makes in order to enhance her current income. Frank must have some benchmark in order to evaluate these decisions, and this is where the literature on happiness comes in.

---

8 See e.g. Golden (2000).
HAPPINESS, INCOME, AND LUXURY GOODS

If the evidence showed that, despite the onset of positional externalities and the zero-sum nature of many consumption decisions, people were increasingly happy, Frank would have a much harder time in questioning the status quo. However, the data does not show increasing happiness (Frank 1999, ch. 5; 2003, 3-6). In developed countries, real income has increased significantly, but based on survey evidence, happiness has not increased. People’s satisfaction adapts based on a rising reference point or a rising aspiration level (Frey and Stutzer 2002, 78). The results of happiness research can be criticized in a variety of ways (Coyne and Boettke 2006), but one big problem is comparing self-reported happiness over long periods of time.

Economists and psychologists have devised many methods to conduct and analyze happiness surveys, but we believe the problem with such evidence is simple and intuitive. If we were to ask people a hundred or a thousand years ago to fill out a happiness survey, it does not seem surprising that we might find similar answers to today, despite the increase in material well-being since that time. The respondent’s notion of what qualifies as “happy” is context specific, which relates to the point about adapting to material conditions. But few would suggest that economic growth did not better the lot of humankind.

If someone had asked Adam Smith to self-report his tallness, he probably would have said that he was average or medium. But today he would be short.

Was George Washington happy? During the course of his life he suffered from diphtheria, malaria, smallpox, tuberculosis, dysentery, quinsy, carbuncles, and pneumonia. From our perspective, much of the time he must have been quite miserable for those reasons, not to speak of his lack of cable TV and other modern components of home entertainment.

Spending on virtually any good could be considered wasteful because the individual is simply going to adapt to having it. And while critics of standard economic assumptions point out the problems with GDP as a

---

9 For recent examples, see Kahneman and Krueger (2006), and Di Tella and MacCulloch (2006).
10 Frey and Stutzer (2002, 87) describe the evidence for people being on a “hedonic treadmill,” in which each person’s ability to experience satisfaction is mainly inherited. They also discuss evidence in which adaptation to circumstances is not as complete as some have claimed, so that absolute conditions matter.

ECON JOURNAL WATCH 418
measure of well-being, most measures of non-material goods that are generally considered to contribute to well-being—such as health, leisure, and the environment—tend to be correlated with GDP (Di Tella and MacCulloch 2006, 32-33).

There are two benchmarks that people use to infuse their context-specific meaning of “happy”: other people’s experience and their own past experience (Frey and Stutzer 2002, 78ff.; Friedman 2005, 81). Frank wants to focus on the comparison to others, but as Benjamin Friedman argues, economic growth provides the potential for everyone to gain in comparison to his or her own experience—and to focus less on relative position (2005, 91). That is why Joel Mokyr suggests that “we should not . . . underrate the capacity of technology to satisfy some of the needs underlying the demand for positional goods” (Mokyr 1990, 303).

**THE INVISIBLE HAND LIVES**

Because of positional externalities, according to Frank, the invisible hand of Adam Smith breaks down:

Far from being a principle that applies in most circumstances, the invisible hand is valid only in the special case in which each individual's rewards are completely independent of the choices made by others. In the rivalrous world we live in, precious few examples spring to mind. (Frank 1999, 271, emphasis added)\(^\text{11}\)

For Hirsch, Adam Smith's “generally benign invisible hand was a favorable inaugural condition of liberal capitalism,” when most consumption was not positional in nature (1976, 11).

\(^{11}\) In finding market failure in such a wide number of transactions, Frank is similar to Joseph Stiglitz, who argues that relaxing the stringent information assumptions of the Arrow-Debreu framework means that the invisible hand is “palsied” at best (Stiglitz 2002, 2). And, similarly, Stiglitz asserts that Arrow-Debreu is “the model that seemingly provides the intellectual foundations of whatever belief one has in the market economy” (1994, 16), ignoring or dismissing the tradition following Adam Smith that does not employ a double standard when considering the reality of government and politics; see e.g. Hayek (1979, 74) on the comparative real-world achievements of free enterprise.
It was not lack of imagination or modern welfare economics that prevented Smith from considering a coercive solution to positional rivalry, but rather the implicit realization that some effects on others are not the proper aim of public policy. Smith was critical of luxury and the vanity of the rich, but his approach was to enlighten people by showing them that their self-interest broadly conceived resided in liberal virtues and the market order. Smith argued that “little else is requisite to carry a state to the highest degree of opulence from the lowest barbarism but peace, easy taxes, and a tolerable administration of justice: all the rest being brought about by the natural course of things.”

When many of these positional arms races are viewed in a standard prisoner’s dilemma framework, it seems that we are stuck in an inferior equilibrium. A collective agreement is necessary, and since it cannot be achieved within the structure of the game as stated, government action as an exogenous factor is necessary. But voluntary agreement can often lead to a spontaneous solution. Efficiency opportunities are not always or necessarily realized, but voluntary incentives (in the broad sense of the term) are more powerful than Frank gives them credit for.

Regarding Hirsch’s tiptoe example, it is noteworthy that we usually don’t see an equilibrium where everyone remains on their feet, although in Frank’s version of it, everyone remained standing at a Diana Ross concert (1999, 152). In many cases, however, norms exist as to proper behavior, and the guy who yells at others to take their seats may not realize his role as an entrepreneur helping to coordinate a superior equilibrium. Other mechanisms might be used by owners with a stake in satisfying their customers. Theaters and sports arenas are usually privately owned, and the owners have an incentive to establish and enforce rules that customers find beneficial on the whole (after all, there are some benefits from getting the audience off their butts). There is of course an extensive literature on how neighbors and even strangers cooperate (a classic work being Ellickson 1991). Research strongly suggests that private ownership conduces to cooperation (Beito et al. 2002).

Frank recognizes the potential for social norms to mitigate positional externalities and conspicuous consumption, but he argues they are inadequate and the government needs to step in (1999, 168-169, 201-203). Larger temptations for defection have caused some of these norms to break

---

Positionality

down (Frank 2004, 133). Unilateral action by some people who want to live a simpler life is not a viable option, according to Frank. “There are many instances in which spending differently would improve matters if everyone did so, yet would make matters worse for any individual who acted alone” (Frank 1999, 190). That may not ring true to many people in middle-class jobs and neighborhoods who are probably unaware of being involved in a positional “arms race” of all against all. The existence of such a middle class would also be evidence against the all-or-nothing implication of winner-take-all markets. Frank emphasizes the data on inequality to support his argument, but the question is whether we view such a trend in inequality as zero-sum or whether incentives for wealth creation among the upper class have some invisible-hand aspect to them.

Frank does not make a convincing case that norms against positional spending have undergone such a significant change—hasn’t the Internet flattened private institutions and fragmented culture?—or that luxury spending inevitably results in an expenditure cascade among all income groups. He opens *Luxury Fever* with an example of his buying a new gas grill; he marveled at high-end models costing $5000 but admits that most grills sold are priced below $700 (1999, 1-2). So what are the implications of such an example? Rich people spend a lot on seemingly frivolous extras, but it is not clear that it has much effect on others; if you have to buy a new grill to impress your friends, you may want to find new friends, or to grin at your friends’ petty vanities. If you are not capable of such solutions, will facing Frank’s high tax rates save you?

Frank writes how he turned down a cheap Porsche because of social norms against conspicuous consumption in a small town, but that since then such norms have broken down (1999, 168-169). It is far from obvious that such a claim can be generalized, and it seems that conspicuous consumption can just as often earn ridicule among your peers instead of esteem or admiration. In our casual estimation, popular films and TV shows are quite effective in disapproving of ostentation, snobbery, and wasteful rivalry. The snob is never the good guy and pretension is made comic. *Seinfeld* and *The Simpsons* are part of the invisible hand—as expounded in *The Theory of Moral Sentiments*.

---

13 According to *The Economist*, middle-class shoppers are “trading up and trading down” at the same time. Consumers are becoming more cost-conscious for some goods, and using some of the savings for buying the occasional luxury. See “The disappearing mid-market,” 2006.
No one claims that markets work perfectly or that all Pareto-improving transactions will be made. Yet, a snapshot view of markets that appear to be winner-take-all or plagued by positional concerns misses some of the ways that people spontaneously adjust. Workers might gravitate to careers in which the importance of relative rankings is consistent with their preferences. An investment banker may feel pressure to wear an expensive suit and consume other conspicuous goods. It is easy to call this socially wasteful spending, but, as natural primps and popinjays, those self-selecting into such a career may aspire to enter this contest, and the contest may have little impact outside of this circle. It is not as if those in investment banking suddenly find themselves in a mutually self-destructive positional arms race and wish for a collective solution so they can all drive perfectly serviceable Hondas and wear sweatshirts to work. And to the extent that there is such an urge, as economists we should recognize how the moral culture tends to evolve so as to approve of and accommodate mutual gains—as witnessed in “casual Fridays” and telecommuting.

THE UPPER CLASS AND ECONOMIC GROWTH

Even critics of a free society are likely to admit, as does Frank (e.g. 1999, 6), that spending on today’s luxuries lead to innovation and higher standards of living for all income groups. The poor in the United States are at the bottom of the status scale, despite their material condition above their ancestors or their counterparts in other countries. Frank’s major point is that positionality is ineluctable. But when the people have means beyond their perceived needs, there is a demand for innovation. As Frank himself notes (2003, 8), Adam Smith was familiar with the type of context-sensitive consumption at issue here. Greeks and Romans, according to Smith (1981, 870), lived comfortably without linen shirts, but now laborers would be ashamed to appear without one; the example closer to Smith’s time was leather shoes. Before Smith, Mandeville wrote of such goods as beer, a plain dress made of cloth, and feather pillows,

[M]any things were once looked upon as the invention of luxury, are now allowed, even to those that are so miserably poor as to become the objects of public charity,
nay counted so necessary, that we think no human creature ought to want them. (Mandeville 1714, 110)

When construction using steel was beginning to see widespread adoption, Andrew Carnegie offered the maxim: “Capitalism is about turning luxuries into necessities” (Johnson 1998, 551). Frank uses these observations as examples of the importance of relative position and comparison with contemporary standards. They are perhaps more usefully considered as examples of the achievements of economic growth.

In the dynamic of a growing economy, the wealthy provide a market for goods that must be expensive in order for supply to be viable. The wealthy pay extra to enjoy the benefits of new goods, which, if suitable to human existence, will later become inexpensive and widely adopted. F.A. Hayek saw this clearly, and even attributed the rapid economic advance that we have come to expect in large measure to inequality: “What today may seem extravagance or even waste, because it is enjoyed by the few and even undreamed of by the masses, is payment for the experimentation with a style of living that will eventually be available to many” (Hayek 1960, 44). Hayek also offered the primary insight about happiness: “The pleasure may be solely in achieving what we have been striving for, and the assured possession may give us little satisfaction” (41). Thanks to economic growth, one’s comforts today exceed those of three years ago.

Of the splendidous houses of the late 19th century, historian Paul Johnson wrote, “the overwhelming majority of the men who built these houses were not leisured; they were fanatical hard workers.” In Frank’s view it would seem such hard work was wasteful. But “the houses they built tended, in both construction and functioning, to advance the cause of high technology,” including such innovations as laundry rooms, ice-freezing rooms, central heating and air conditioning (Johnson 1998, 593). Today, Bill Gates’s home uses technology that has not yet reached the upper-class market, let alone the masses. Government advocates are often ready to propose that the government spend taxpayers’ money to advance science and technology. But what better way is there to advance technology for human existence than for wealthy entrepreneurs to experiment with their own money in building laundry rooms, ice-freezing, and central heating that they themselves have to live with?

Frank argues that higher consumption taxes would reduce positional externalities and encourage savings, thereby increasing growth. So for him there is no necessary tradeoff to be made. Frank argues that, living under the proposed consumption tax, top earners would have to make only trivial
sacrifices. Frank suggests that they will have to content themselves with a Ferrari rather than tomorrow’s supercar (Frank, 1999, 221). The rich will still be paying a status premium, but what some of these examples miss are goods that haven’t been invented yet, or that haven’t yet become mass-produced and will eventually be considered commonplace. Moving beyond yesteryear’s Ferrari is part of that process.

THE HOUR LIVED: CHOOSING WORK AND LEISURE

A large part of Frank’s argument relies on the pursuit of luxuries and positional goods such as advantaged residence at the expense of leisure activities that would have more happiness “bang” per hour lived. However, some evidence indicates that the decline in leisure may largely be a myth. Aguiar and Hurst (2006) find that leisure increased significantly between 1965 and 2003. The conclusion is robust to various measures of leisure. While market hours of work were relatively constant over the period, average leisure per week increased by about 6 to 8 hours for men, and 4 to 8 hours for women. Yet Frank buys into the false impression. The Economist suggests that the false impression arises from rising incomes, which make time worth more, improving technology, which make people available on short notice even when not working, and increasing competition, which makes people less secure in their jobs (“The land of leisure,” 2006).

The amount of leisure taken by most workers in the U.S. is often compared unfavorably to Europe. Alesina et al. (2005) conclude that the fewer hours worked in Europe result from government regulation, including union privileges. The authors refrain from any conclusion about whether the policies decrease or increase welfare. The increase would come by coordinating what Frank would call a positional arms control “agreement”—“I won’t earn much during the summer if you don’t”—even though this arrangement comes by way of government force.

14 Frank notes the debate over trends in leisure and work hours (1999, 48-51).
HOW MUCH LESS POSITIONAL ARE “LEISURE” ACTIVITIES?

Frank’s big plan is to lay more taxes on the cash nexus—either at receiving income or expenditure—because many of the things we obtain with money are positional. The whole scheme implies that the cash-oriented activities are significantly more positional than the non-cash-oriented activities, which are left untaxed. If, instead, the two realms of activities were in fact equally positional—if positionality in fact characterized human pursuits quite universally—then Frank’s plan would make no sense.

Suppose an economist decided that human beings were causing global warming by exhaling carbon dioxide, and proposed that taxes be increased, on the reasoning that, since people breathe while they work for the money, we should tax the cash nexus. One of the big flaws in such a proposal is that, although the new taxes would shrink the cash nexus, they would not reduce breathing, since during time spent “off the job” people do not breathe any less. Even if we grant that the carbon dioxide in this story or that the positional striving in Frank’s story is a negative externality, we should be mindful of the fundamental issue of whether the tax-discouraged activities are more externality intensive than the untaxed (or less taxed) activities. Maybe Frank can make a good case that the cash-oriented activities are more positional than the non-cash activities, but the major arguments given for this assumption are thought experiments about vacation time, commuting time and other hypothetical tradeoffs (1999, 79-90; 2005, 137). In Frank’s analysis, what really matters in the end is hours of life experienced, and whether positionality is being pursued with them. He presumes that cash activities are especially positional. Should we accept that presumption? What about cash expenditure that is not especially positional? Is the positionality of cash-oriented hours lived obviously greater than of non-cash hours? In their “leisure” people may be educating their children to compete in the positional “rat race,” improving their golf game, getting a sun tan, working out to look better than the competition, or reading books to be more clever and worldly than their associates.

Moreover, the cost of a positional good consists of the price paid to the seller and transaction and shopping costs. To acquire a gas grill, a sports car, or a house in a good school district, one typically visits vendors, discovers the offerings, researches the characteristics, compares quality, gets
assurances, wrestles with the decision, negotiates and scrutinizes the deal, attends to delivery, and learns how to use, enjoy, and flaunt his new positional good. If Frank raises income-tax rates, people will to some extent substitute into these “leisure-time” inputs that remain untaxed, with the result that hours devoted to positional climbing have declined less than Frank might have expected. People work less and shop more carefully. If Frank’s imposition takes the form of consumption taxes, then people might substitute into tax evasion and avoidance.

Frank places great emphasis on vacation as an example of a good that is nonpositional. He even uses it to illustrate the meaning of positionality (Frank 2005, 137). However, as a kind of “anti-work,” vacationing may be tied up with positionality. One reason you might not be embarrassed at having two weeks less vacation than everyone else per year (as in Frank’s thought experiment) is that having less vacation time can signal that you are too important at your workplace to disappear for weeks at a time. Also, long spells of vacation can signal that you do not like your work—a condition that no one can envy. Taking little vacation can signal that you love your work, which is probably the condition most correlated to true happiness. The upshot is that it is not easy to draw a line between activities that are positional and activities that are nonpositional. In terms of positional goals, activities inter-relate as complements and substitutes, and who knows how people will respond after “the man of system” tries to rearrange “the great chess-board of human society” (Smith 1982, 233-234).

Again, maybe Frank can make a case that positionality waste is more pronounced in cash-oriented activities, to be subjected to proposed new taxes, than non-cash activities, but the case is yet to be made. Adam Smith strove to establish a presumption of liberty, to place the burden of proof on the coercionists. Frank certainly has not lived up to the Smithian incumbancy.

EXCELLENCE IS ABOUT EXCELLING

In The Theory of Moral Sentiments, Adam Smith says that if “a man of humanity in Europe” had to sacrifice his pinky to prevent a terrible earthquake in the far-off land of China, he would do so. Smith explains:

It is not the love of our neighbour, it is not the love of mankind, which upon many occasions prompts us to the
practice of those divine virtues. It is a stronger love, a more powerful affection, which generally takes place upon such occasions; the love of what is honourable and noble, of the grandeur, and dignity, and superiority of our own characters. (Smith 1982, 137)

According to Smith, cultivating our duty to humanity serves our sense of superiority. Although our moral acts might be “seen” only by the impartial spectator, that judge flows in Smith’s theory directly from actual observation and social propriety, where people witness and esteem moral action. Would Robert Frank contend that magnanimity is a positional arms race?

To clarify the meaning of positionality, Frank (2005, 137) gives a thought experiment: “[Y]ou must choose between two worlds that are identical in every respect except one. The . . . choice is between world A, in which you will live in a 4,000-square-foot house and others will live in 6,000-square-foot houses; and world B, in which you live in a 3,000-square-foot house, others in 2,000-square-foot houses.” Frank reports (without citing any evidence) that “most people say they would pick B.” We are skeptical. Not only does choice A afford one an additional 1,000 square feet, but it serves one’s concern for the grandeur, dignity, and superiority of his or her own character, particularly if the choice is made public. Indeed, to choose B, putting everyone else in 2,000-square-foot homes rather than 6,000 just so you can be on top, is downright malicious. Were this type of choice to be a genuine human regularity, it is likely that voluntary institutions would evolve to make the choice subject to approbation and disapprobation and arrive at the better outcome.

Just about every form of excellence, every aspiration, can be said to contain an element of positionality. Striving for better skill in hockey, horseback riding, or hopscotch is a positional arms race. In Smith’s theory of moral evolution, practically every form of human meaning and significance comes by way of the “looking glass” that is the societal reaction to our conduct.

Desires that do not involve positionality would perhaps be comforts and sheerly sensuous pleasures that do not entail any kind of striving. Conformity to routine and stasis would seem to be the activities that best avoid positionality. Frank wants us all to take longer vacations, but what if, rather than traipsing through tourist destinations, lying around in the sun, or reading unexceptional novels, we instead take the opportunity to retool and learn, in striving to be better than we are, and, perforce, better in relation to others?
To judge whether individual instances of positionality are pernicious or ameliorable, we need to develop some sense of positionality as it manifests itself in the wide array of human activities. Here, again, we feel that Frank’s treatment is unsatisfactory. He argues as though it is sufficient to identify some element of positionality in the activities he proposes to tax.

**HOW MUCH LESS POSITIONAL ARE GOVERNMENTAL ACTIVITIES?**

Since Frank never speaks of making the tax proposal revenue-neutral, it presumably would have the consequence of transferring additional resources to the government. Implementing Frank’s progressive consumption tax “will free up literally trillions of dollars each year to spend in ways that will create lasting improvements in the quality of our lives” (1999, 279).

By scaling back on the rate at which luxury consumption has been growing we could drink safer water, breathe cleaner air, and eat food that is less likely to make us seriously ill. We could have more autonomy, variety, and safety in the workplace. We could fill the potholes in our streets, fix our crumbling bridges, and provide more help for people with serious drug problems. We could save enough to retire with complete financial security. (Frank 1999, 107)

Once again, Frank makes a basic assumption without even pausing to point out that he is doing so. Government—the “we” filling the potholes etc.—supposedly avoids the positionality problem.

Consider again the analogy of taxing breathing. If the tax revenue is going to augment other human activity that involves just as much breathing, then the policy is obviously getting nowhere in terms of its stated objectives, not to mention the collateral losses it entails.

Maybe Frank could make a case that political society is less positional than civil society, but he has not even attempted to do so. It is by no means obvious that political society is less positional. Frank’s proposal would surely expand government employment. At George Mason University, which
Positionality

is situated in northern Virginia near Washington, D.C., the Masters level economics classes are populated by many federal government employees who are looking to move up. The GS-5s want to become GS-6s, and so on. They need a Masters degree to do so, so the Masters level classes are filled with students who often have little real interest in the subject matter. The promotion is not merely an increase in pay but a gain in position and prestige. Much of federal government employment is organized in this fashion, like one great military organization. The emphasis on rank and position seems to us to be more uniform, conspicuous, and pronounced in the government sector than in the private sector.

Besides the matter of employment rank, we may ask about other positional contests in the government arena. Politicians compete for fame and celebrity, administrators compete for rank, influence, and power, and grant recipients compete for prestige and honors. People compete for the favor of the powerful, and not just for material purposes. It may be a genetic tendency to strive for the favor of the group’s leader or “alpha male.”

Again, our point is not that government is more positional than civil society. It is merely to point out that Frank has been remiss in not even bothering to justify his crucial assumption that it is less so.

Other Realities about Government

Frank might say that positional goods are zero-sum. But much of what happens in politics and government is negative-sum. Instead of politicians competing to provide a needed public good, competition in the political sphere can get us a statue of Robert Byrd in the West Virginia Capitol, the infamous “bridge to nowhere” in Alaska in 2005 (Grunwald 2006), rail transit systems that few people use, the public school system, and mass incarceration for consensual crimes. Such political externalities, which are more or less inherent in majoritarian politics, have far less scope for corrective mechanisms than in the voluntary sector. Frank does not seem to be impressed with theories of resource utilization that maintain that

15 Byrd was the first politician to obtain more than $1 billion in pork for his state, according to the Citizens Against Government Waste, www.cagw.org.
efficiency usually flows best from private ownership, residual claimancy, and voluntary choice.

Taxes and Distortions

Despite research purporting to show otherwise (e.g. Lindert 2004), there is still significant evidence of distortions and deadweight losses from taxes. Davis and Henrekson (2004) find that a tax rate difference of 12.8 percentage points (one standard deviation across countries) leads to 122 fewer market work hours per adult per year, a drop of 4.9 percentage points in the employment-population ratio, and a rise in the shadow economy equaling 3.8 percent of GDP. Frank would like to use taxes specifically in order to reduce hours of employment, but the disemployment in Europe is probably welfare-decreasing.

CONCLUSION

Robert Frank has greatly illuminated the importance of relative position. However, he has taken his insights in nefarious directions—namely, justifying institutionalized coercion. As Stiglitz can for information problems, Frank can find externalities in a wide range of market transactions, and any externality leads to the theoretical possibility of a beneficial government intervention. But transactions costs in the real world prevent many of these exchanges, and there should be no presumption that government can do any better, particularly when we are talking about curtailing a preference that is inherent in human nature. In this case, the relevant transactions costs should also include the effects of a significant increase in taxation to curtail the positional “arms race.”

The plausibility of Frank’s arguments for extensive market failure requires various assumptions, including the usefulness of happiness comparisons over time, the widespread existence of winner-take-all markets, the failure of voluntary evolution to internalize and mitigate externalities, and significantly less positionality in both “leisure” and governmental activities than in the full range of activities Frank proposes to tax.

Throughout history, observers have in one form or another lamented the positional “arms race” and the seemingly wasteful spending on luxury
goods. But one era’s luxuries are another’s commonplace goods. A significant increase in tax revenue in order to transfer resources from luxury spending to government spending could lead to the disincentive effects and waste that high consumption taxes are attempting to avoid. It’s not clear what kind of society Frank envisions, but it seems to be one where policy tries to dampen the private striving for eminence and expand the role of government. To the extent that we do climb over each other in the contest for positional goods, the disease may be less bad than the supposed cure. Buying expensive gas grills is a relatively harmless way for people to flatter the will to eminence.

REFERENCES


The disappearing mid-market. The Economist, May 20, 2006, p. 68.


Hofstadter, R. 1955. The Age of Reform; From Bryan to F.D.R. New York: Knopf.


ABOUT THE AUTHORS

Andrew Kashdan is a Ph.D. student at George Mason University. His email address is akashdan@gmu.edu.

Daniel Klein is professor of economics at George Mason University, and associate fellow at the Ratio Institute in Stockholm, Sweden. He is the chief editor of Econ Journal Watch. His email address is dklein@gmu.edu.
Taking Libertarian Concerns Seriously: 
Reply to Kashdan and Klein

ROBERT H. FRANK*

I WANT TO BEGIN BY OFFERING MY SINCERE THANKS TO Andrew Kashdan and Daniel Klein for their spirited critique of my work. They have read me carefully and raised many important questions. But before responding to them in detail, I want to say something about what they perceive as my shortage of “libertarian sensibilities about politics, government, and society.”

This charge has been leveled at me by other readers of my work. One of the earliest reviews of Luxury Fever posted on Amazon.com, for example, said simply, “Good morning, comrade! All the world needs is one more socialist with a Ph.D.” It is a frustrating charge because it is one that no one who knows me well would make. As my friends can attest, I would place near the top of any reasonable scale of libertarian sensibilities.

A profound reluctance to surrender personal autonomy is perhaps the most widely shared personality trait among libertarians. This trait is well represented among academics. Indeed, many of us became professors precisely because it is one of the few jobs where we work for no one and no one works for us.

But no matter how strong their preference for personal autonomy, libertarians with even a modest capacity for reason are forced to retreat from the most extreme versions of libertarianism. Some versions, for example, insist that all taxation is theft. But because of the free-rider

* Johnson School of Management, Cornell University. Ithaca, NY 14853.
I thank Tom Gilovich and Ori Heffetz for helpful discussions on an earlier draft.
problem, a nation in which tax payments were purely voluntary would be unable to field an army. It would eventually be invaded by some other country’s army, and its citizens would then be forced to pay taxes to the government of that country. So even people with the most extreme libertarian sensibilities generally concede the necessity of empowering government to collect taxes by force of law. They don’t like taxes but recognize that the alternative is clearly worse.

Taxes are not the only issue on which most libertarians have felt the need to compromise. They must also make reasoned judgments about the conditions under which the state can restrict the individual’s freedom of action in other ways. And most are prepared, for example, to concede the state’s right to prevent them from driving 90 miles per hour in school zones.

Yet there remain a small number of extreme libertarians who reject all compromise. They denounce all taxation as theft and reject any attempt by the state to restrict their behavior as an infringement on their individual rights to do as they please. These libertarians are really just crybabies. Having long since forfeited any claim to be taken seriously in debate, they wield no power in defense of liberty.

Professors Kashdan and Klein are not extreme libertarians, and neither am I. Each of us accepts that the state may tax us and enforce various laws that prevent us from doing as we please. So where do we differ?

My own account of libertarianism holds that the state can restrict someone’s freedom of action only if it can be shown that failure to do so will result in unacceptable harm to others. Kashdan and Klein counter that libertarianism is rooted in property ownership and freedom of contract, “not guarantees against any kind of harm.” Well, yes, but the two statements are completely compatible. Kashdan and Klein are right, for example, that property ownership and freedom of contract permit business owners to inflict extreme economic hardship on competitors. But the laws of property permit this only because preventing owners from competing vigorously would cause even greater harm to them. In my account of libertarianism, preventing harm to others is a necessary, but not sufficient, condition for restricting behavior.

Kashdan and Klein do not dispute that positional arms races can lead to welfare losses for reasons precisely analogous to those we see in the case of military arms races. In each instance, the necessary condition for welfare losses is that context, or relative position, matter more for some expenditure categories than others. Too much is spent on military arms, for example,
because relative position matters more for bombs than for most consumption goods. Likewise, too much is spent on positional goods because context matters more for such goods than for nonpositional goods. Of course, the mere fact that an allocation is imperfect is hardly sufficient to justify restricting individual behavior. Perhaps, for example, we have little reliable information on the extent to which evaluations depend on context in different domains, in which case we would not know which goods to promote with collective action and which to inhibit.

Certainly, it would be fair to describe research on how context shapes evaluation as being in its infancy. Yet we are almost always forced to act on the basis of less than complete information. Conventional economic models entail a very specific, if implicit, assumption about the extent to which context sensitivity differs across domains: they assume that context doesn’t matter at all. It is thus little wonder that researchers who employ these models are reluctant to state this assumption explicitly. It is one that cannot withstand even minimal scrutiny, for although we do not know the precise extent to which context influences evaluation in different domains, evidence suggests not only that it matters, but also that it matters much more in some domains than in others. There is persuasive evidence, for example, for the proposition that safety is less positional than other consumption categories as a whole.1 I know of no evidence at all supporting the contrary proposition. The claim that positional concerns in unregulated labor markets will produce an equilibrium in which workers choose inefficiently low levels of safety is thus completely consistent with all available evidence.

But as Kashdan and Klein are correct to point out, even this conclusion does not ensure that we would do better by enacting safety legislation. After all, even the most well-intentioned laws and regulations sometimes do more harm than good. For some reason, Kashdan and Klein consider me insufficiently attentive to this possibility. As my students can attest, however, I take great pains to emphasize it, not just in my writings, but also in my courses. Merely discovering that an unregulated market allocation is inefficient, I tell them, provides no assurance that a regulatory solution will be better. I stress this point repeatedly. In the safety domain, I often illustrate it with examples like the following passage on requirements for ladders, taken verbatim from the U. S. Occupational Safety and Health Administration’s 1976 manual of workplace safety standards:

---

1 For a survey of this evidence, see Frank (2007, chapter 7).
The general slope of grain in flat steps of minimum dimension shall not be steeper than 1 in 12, except that for ladders under 10 feet in length the slope shall not be steeper than 1 in 10. The slope of grain in areas of local deviation shall not be steeper than 1 in 12 or 1 in 10 as specified above. For all ladders, cross grain not steeper than 1 in 10 are permitted in lieu of 1 in 12, provided the size is increased to afford at least 15 percent greater strength than for ladders built to minimum dimensions. Local deviations of grain associated with otherwise permissible irregularities are permitted. (Quoted by R. Smith 1977, 11-12)

This befogged passage appeared in a section of the OSHA manual devoted to ladders that is 30 pages long, two columns to the page. It is easy to imagine that after having read this passage, many managers decided to eschew any attempt to master these regulations and instead simply abandoned any activities that required ladders. That step, needless to say, would have been unlikely to make the workplace any safer.

In short, I am happy to stipulate that the existence of a market imperfection does not guarantee that regulation will improve matters. But neither does the mere possibility of counterproductive regulation constitute a persuasive case against all forms of regulation. In the environmental domain, for example, we have discovered that requiring marketable effluent permits often reduces cleanup costs to only a fraction of what they were under heavy-handed, prescriptive regulation. Similarly, it may be that injury taxes would be both more efficient and less coercive than many forms of prescriptive safety regulation. But these are empirical questions, not questions of ideology.

Kashdan and Klein also emphasize, and I agree, that private solutions to externalities are often sufficiently compelling to render governmental intervention unattractive. Here they follow the trail blazed by Ronald Coase, who pointed out that if transaction costs were sufficiently low, private parties could negotiate efficient solutions to problems caused by externalities. As Coase emphasized, the efficient solution in general is to place the burden of adjusting to an externality on the party who can do so at the lowest cost. In some cases, for example, it might be more efficient to install a filter on a factory’s smokestack, while in others it might be cheaper for downwind residents simply to move.
It is this line of thinking that has provoked many of the most vehement objections against legislative action to curb positional externalities. Such objections typically come from those who believe that positional concerns are rooted in negative emotions like envy and jealousy, and are hence an illegitimate basis for restricting anyone’s behavior. In their eyes, shaping public policy on the basis of positional concerns would be like giving policy weight to the preferences of sadists.

This objection has considerable rhetorical force, for society has a strong interest in discouraging envy and jealousy. Indeed, it would be difficult to find anyone who would endorse preventing someone from building a larger house merely because it might make his neighbors feel bad. But as Kashdan and Klein acknowledge, I have taken pains to emphasize that the cost of having inferior relative position typically has nothing to do with envy.

The most vivid example involves the link between relative position and the cost of providing a good education for one’s children. In the United States, as in most other countries, the perceived quality of local schools closely tracks local real estate prices. If you want your child to attend a public school of average quality, you must therefore spend roughly the median price for houses in your area. But a “good” school is an inherently positional concept. It is simply one that compares favorably with the relevant local alternatives. And because school quality is defined in relative terms, the logic of musical chairs implies that half of all students must end up in bottom-half schools, no matter how much everyone spends.

The median family on the income scale rationally believes it can do better by increasing the amount it spends on housing. But when all families follow suit, the effect is merely for everyone to have spent more on housing. No one moves forward in relative terms. And if their bidding is financed in part by sacrificing nonpositional items such as safety, leisure, and savings, the result is inefficient.

Admonishing people to ignore their feelings of envy is a productive step, to be sure. But one cannot realistically expect parents to abandon their goal of sending their children to the best schools they can afford. If mobility were completely costless, it might be possible for them to persuade like-minded others to form new communities with modest houses and generous school budgets. But given that most people also value living in reasonably close proximity to where they work, forging a private solution of this sort would be difficult at best. And so the practical choice confronting families almost always entails a dilemma: buy an expensive house in a
neighborhood with good schools, or a cheaper house in a neighborhood with bad schools.

Like expenditures for schooling, expenditures for a host of other goods entail positional components that have nothing to do with envy. When a couple goes out to dinner for their anniversary inducing envy in their friends and neighbors is almost certainly not their aim. Rather, they just want to share a memorable meal. But a memorable meal is a quintessentially relative concept. It is one that stands out from other meals. Thus, beyond some point, additional expenditures do not augment the number of memorable meals.

It is the same with cars, which are judged according to how their characteristics compare with those of other cars in the same local environment. The same car that would have been experienced as having brisk acceleration and sure handling by drivers in 1950, for example, would be much less charitably evaluated by most drivers today.

There are of course some people who buy things in the hope of making others feel envious. And there are others who buy things in order to assuage their own feelings of envy. But examples of both types are relatively rare. In portraying positional arms races as entailing purchases primarily driven by such motives, Kashdan and Klein consign them to the periphery. Such arms races, they conclude, are not only unimportant, but also not legitimate grounds for policy intervention. But positional externalities are really just context externalities. And in varying degrees, context matters everywhere.

Given their professed libertarian sensibilities, I am puzzled that Kashdan and Klein take such an uncharitable view of my proposal to use tax policy for reducing the costs that arise from positional arms races. My own libertarian sentiments dictate that if the goal is to alter our behavior for common good, the instrument we choose should afford people as much flexibility as possible. Compared to the alternative of prohibiting specific behaviors, it is generally far less coercive simply to tax them.

This, as noted, is the central lesson of our experience with policies aimed at reducing environmental pollution. Early policies required across-the-board reductions in effluent discharges, an approach that proved grossly inefficient, because it made no allowance for the fact that it costs some firms far more than others to curtail pollution. Effluent taxes or marketable permits generally work much better, because they concentrate pollution reduction in the hands of those who can accomplish it at the lowest cost.

The same principle is directly relevant for changing behavior in other domains. Consider, for example, the laws in many states that require cyclists
to wear helmets. These laws have been defended on grounds that injured cyclists often impose heavy costs on others. But that goal could be achieved more directly by the less intrusive step of requiring cyclists to carry adequate insurance. The actual reason for helmet laws is that legislators believe they should prevent cyclists from taking the irrational risk of riding bareheaded.

New York State requires bicyclists 16 and under to wear helmets, a law that almost surely saved the life of one of my sons, who was involved in a serious bicycle accident when he was 13. Such laws enjoy broad support. But many libertarians oppose similar laws for adults, as I once did, in the belief that people of voting age should be free to make such decisions for themselves. With additional exposure to the relevant social science literature, however, I have become persuaded that many adults are no more rational than most children when deciding whether to wear a helmet. And even fully rational adults may experience collective action problems when confronting this decision. They might be concerned, for example, that a voluntary decision to wear a helmet would signal fearfulness or a poorly developed sense of fashion.

Suppose, for the sake of discussion, that for these or other reasons, the vast majority of adult cyclists favored a law requiring helmets. Libertarians would still have intelligible reasons for objecting, for there are at least some cyclists who insist they feel deprived of the essence of a cherished experience if they are forced to wear helmets. For these people, riding without them is probably rational. To require them to wear helmets just because others might choose irrationally has always seemed illegitimate to me.

But requiring helmets for all is not the only policy option. Alternatively, we could permit cyclists who feel strongly about riding bareheaded to exempt themselves from the requirement by purchasing medallions that could be displayed on their license plates. If the fee were, say, $300 a year, most of the people for whom riding bareheaded was in fact an irrational decision would start wearing helmets. And those who felt strongly about riding without one would retain the option to do so. Revenues from the sale of medallions could be used to offset other taxes.

Taxation as an instrument for curbing positional arms races is attractive for exactly parallel reasons. Rather than outlaw socially costly forms of consumption, we can achieve the same result more cheaply and less coercively simply by taxing them. Kashdan and Klein complain that taxation is a blunt instrument because it “brings force against the guilty and innocent alike.” Would they offer a similar complaint about effluent
charges? The optimal effluent tax depends on the aggregate harm caused by the pollutant in question, not on the motives of polluters. It would make no sense to exempt a specific firm from an effluent tax on the grounds that it had no intention of harming others. Similarly, it would make no sense to exempt a consumer from a consumption tax merely because it was not his intention to harm his neighbor. Positional externalities impose costs on others, irrespective of people’s intentions. If feelings of envy were the only such cost, the best policy response might be simply to exhort people to suppress those feelings. But, as noted, the primary costs associated with positional externalities have nothing to do with envy. And those costs are not so easily avoided.

The progressive consumption tax that I have proposed is relatively simple. Taxpayers would report their incomes to the IRS just as they do now. They would also report how much they had saved during the year, much as they do now to exempt money saved in IRAs and 401(k) accounts. People would then pay tax on their “taxable consumption,” which is just the difference between their income and their annual savings, less a standard deduction. Rates at the margin would rise with taxable consumption.

If the tax were revenue neutral, marginal rates at the top would be significantly higher than current marginal tax rates on income. Such a tax would thus alter the incentives confronting wealthy taxpayers as they decide how much to spend and save. One direct effect of this change in tax policy, then, would be to reduce current expenditures on mansions. But the policy change would also have indirect effects. If housing is, as evidence suggests, a strongly positional good, the fact that people at the top save more and spend less on it will shift the frame of reference that influences the housing expenditures of those just below the top. So they, too, will spend less on housing, and so on all the way down the income ladder.

By all available evidence, that would be a good thing. In 2005, the aggregate household savings rate in the United States was negative. Americans spent more than they earned during a full calendar year for the first time since The Great Depression. Liberals and conservatives alike agree that our failure to save has had damaging macroeconomic consequences, that we would all be better off if we all spent less and saved and invested more. But no individual has the power to alter the aggregate savings rate, which shares many characteristics of traditional public goods.

Perhaps Kashdan and Klein believe that our current lack of savings rates is not a proper object of social concern. If so, it would be interesting to hear why. But if, more plausibly, they believe that saving more would be better, on what grounds might they object to my proposed switch to a
progressive consumption tax? One could always trot out the “taxation as theft” objection, but as noted earlier, that objection has been long since abandoned by serious libertarians. Given that we must tax something, what should we tax?

By taxing income, we tax both consumption and savings. But given that we save far too little, what possible sense does it make to be taxing savings? In a paper published in 1943—another time when Americans needed to save more—Milton Friedman suggested the progressive consumption tax as the best instrument for raising the revenues needed to pay for the war effort. His argument in that paper is just as compelling now as it was then.

Kashdan and Klein are also critical of the happiness literature I cite. But I have criticized this literature in very similar ways (Frank, 2005). The two central findings reported by happiness researchers are that 1) beyond some point, when everyone’s incomes grow in tandem in a given country, measured happiness levels do not rise; and 2) in a given country at a given point in time, high-income people are happier, on average, than low-income people. Many conclude from the first finding that economic growth is not welfare-enhancing. But as I have argued elsewhere, this conclusion is highly suspect (Frank, 2005). Existing happiness measures clearly fail to capture much that we value. Even so, these measures are strongly linked to a variety of objective indicators of human well-being. And not even the fiercest critics of the happiness literature question the importance it assigns to relative income (see, for example, Kahneman et al., 2006).

The positional arms race model does not require that utility be independent of absolute consumption levels. As discussed earlier, the critical assumption is that context matters more for the evaluation of some goods than of others. The fact that existing happiness measures fail to capture all important aspects of human welfare does not imply that context does not matter, or that its importance does not differ across domains.

Kashdan and Klein also question my claim that much high-end consumption is wasteful. Thus, they write, “if the standard for luxury spending has risen, so that the rich must buy ever more expensive cars to achieve elite status, the amount of waste may not be very significant . . . [because] much of the premium they pay ends up as producer surplus.” This claim rests on the implicit assumption that producers need not compete for the patronage of high-end consumers. This seems a strange position indeed for free-market economists like Kashdan and Klein to defend. After all, the most important difference between the wealthy and others is that the wealthy have more money. Like everyone else, they want good value for their outlays. They
may be willing, even eager, to pay top dollar; but in return, they want the best products available. And improvements in quality cost real resources.

As the upper echelons of the relevant quality distributions are approached, the cost gradients often rise steeply. For example, the Porsche 911 Turbo, which accelerates from zero to sixty in 3.9 seconds, sells for about $150,000. Until recently, it was difficult to find another car with clearly better handling and performance. Then Porsche’s own Carrera GT, introduced in 2004, raised the bar slightly. In the zero-to-60 sprint, it beats the Turbo by two-tenths of a second, and it is slightly more sure-footed on the track. To get these improvements, however, Carrera GT buyers must shell out nearly three times as much as for the Turbo. Some of that money is producer surplus, but only a small fraction. The lion’s share of the company’s revenue is paid out to its factors of production. It is a fundamental misunderstanding of the competitive process to believe that high-end consumption dollars end up largely as producer surplus.

Kashdan and Klein also argue that the waste of luxury spending is circumscribed by the ability of sensible people to choose their friends and co-workers. They note, for example, that “An investment banker may feel pressure to wear an expensive suit and consume other conspicuous goods. It is easy to call this socially wasteful spending, but, as natural prims and popinjays, those self-selecting into such a career may aspire to enter this contest, and the contest may have little impact outside this circle.”

All that may be true, but how does it speak to the question of whether such spending is wasteful? I stress again that by wasteful I mean inefficient from the collective, as opposed to the individual, vantage point. Many high-end purchases are wasteful in this sense because they are driven largely by their functions as signals—both of ability and of the importance of specific relationships. First impressions are often important in competitive environments, for example, and are often influenced in subtle but powerful ways by the kinds of clothing or jewelry people wear. Thus, as a Madison Avenue retailer described a conversation he overheard between two men, each wearing a five-figure wristwatch, “It’s all about who has what. The friend sees his friend has a [Patek Philippe] Pagoda, and these are people who have a certain intuitiveness; they know how much things cost. They ascertain what a guy’s capability or monetary status is by looking at his watch. They know if he’s a player. Or they think they know” (Kuczynski 1998, 3).

A wealthy businessman recently spent $10 million on a coming-of-age party for his daughter. He was undoubtedly motivated in part by a
sincere wish to make her feel special. It was probably not his intention to raise the bar that defines what counts as special.

Most of us will of course never spend millions on a party for a loved one or own a Patek Philippe watch. Yet such spending by top earners has spillover effects. Ooggles-N-Googles, a company described as an over-the-top event planner for children’s parties, has been in such demand that it recently announced that it would begin licensing franchises. Few middle-income families purchase the services of this company. But median expenditures on children’s parties and on gifts more generally have been growing rapidly. Even in middle-income families, for example, a $200 wristwatch is no longer viewed as a suitable gift to mark a spouse’s promotion. When signal strength and subjective evaluations depend heavily on relative expenditure, little would be sacrificed if all spent less on expensive wristwatches and birthday gifts.

Kashdan and Klein write as if such waste were a matter of social indifference. This is a strange belief indeed for economists to hold, for if we had additional resources, there would be innumerable ways in which we could spend them to mutual advantage. For example, budget deficits have compromised our national security, as when the Bush administration reduced financing by 8 percent in 2004 for the Energy Department’s program to secure loosely guarded nuclear stockpiles in the former Soviet Union. Sam Nunn, the former United States senator, now heads a private foundation whose mission is to raise private donations to expedite this effort. And despite the rational fear that terrorists may try to detonate a nuclear bomb in an American city, most cargo containers still enter our ports without inspection.

Budget deficits have also led to cuts in federal financing for basic scientific research, even as our comparative advantage in the scientific domain has continued to dwindle. Describing these cuts as a threat to our long-term economic prosperity, Senator Pete Domenici, Republican of New Mexico, said: “We thought we’d keep the high-end jobs, and others would take the low-end jobs. We’re now on track to a second-rate economy and a second-rate country.”

Citing revenue shortfalls, we have postponed maintenance of our streets and highways, even though doing so means having to spend two to five times as much on repairs in the long run. In the short run, bad roads cause thousands of accidents each year, many of them fatal, and an annual average of $120 in pothole damages per vehicle.

Revenue shortfalls and low household savings rates have also forced us to borrow hundreds of billions of dollars each year, primarily from
China, Japan, and South Korea. These loans, which must be repaid in full, with interest, create additional risk of international monetary instability in the meantime.

The revenue shortfalls that have precipitated these spending cuts are in large measure a consequence of tax reductions enacted during George W. Bush’s administration. Those tax cuts, which put more money in my pocket, were marketed with libertarian slogans. Representative Tom Osborne of Nebraska, for example, defended them on the House floor by saying, “The bottom line is that it’s your money, and you know how to spend it much better than anyone in Washington, D.C.” As a libertarian, however, I cannot see how the resulting public and private expenditure shifts have increased my personal autonomy.

Kashdan and Klein voice concern that to increase funding for government activities would create positional waste of its own. They cite unproductive arms races among civil servants seeking advanced degrees, thereby to increase their salaries under government pay formulas. Having worked in government, I have witnessed government waste at close range. But wouldn’t it be better to attack such waste directly, rather than by cutting funds for both wasteful and efficient programs alike?

Kashdan and Klein also accuse me of being insufficiently sensitive to the positive externalities generated by high-end consumption. Clearly they are correct that many of the technological innovations that were originally developed for the luxury market find their way over time into products bought by others. And such spillovers are clearly a good thing. But on what grounds would it be reasonable to expect that a progressive consumption tax would stifle innovation on balance?

In the short run, the tax would not change the total level of spending. Rather, it would shift the composition of spending in favor of investment. Innovation is hardly confined to the consumption sector. Producers of capital goods also have strong incentives to come up with useful innovations. And with the greater aggregate investment spending caused by a consumption tax, more resources than before would be available for research and development. There is thus no reason to expect innovation to slow down, even in the short run.

In the long run, which is what really counts for the point Kashdan and Klein are attempting to make, their argument collapses completely. Higher rates of investment mean a higher rate of income growth, which means that consumption along the high-savings trajectory will eventually exceed what it would have been had we remained on the low-savings trajectory. From that point forward, there would be more expenditure on
innovation in both the consumption and capital goods sectors.

Kashdan and Klein also question my claim that leisure is broadly nonpositional. Certainly there could be circumstances in which this claim would be false. For instance, in societies in which the greatest concentrations of wealth resulted not from earnings but from inheritance, leisure could even be highly positional—a relatively pleasant way of signaling one’s material standing. But when income and wealth depend largely on personal effort, compelling evidence suggests that leisure will be highly nonpositional. Historically, position in human societies has always been strongly associated with control over material resources. In a famine, for instance, whether a person got fed depended on relative wealth, not relative leisure. Similarly, in a polygynous society, whether a man was able to marry depended on his relative wealth, not his relative leisure. And although most societies are no longer polygynous, relative wealth levels remain an important determinant of sexual attractiveness.

Elsewhere, I have motivated the claim that leisure is nonpositional with a pair of thought experiments. In the first, most respondents are willing to accept smaller absolute house size in return for larger relative house size. In the second, most respondents are willing to sacrifice relative advantage on the leisure scale in order to obtain absolute advantage on that scale.

Kashdan and Klein offer what they apparently believe to be a competing interpretation of the latter choice pattern. Thus, they write that “One reason you might not be embarrassed at having two weeks less vacation than everyone else is . . . that having less vacation can signal that you are too important at your workplace to disappear for weeks at a time. Also, long spells of vacation can signal that you do not like your work—a condition that no one can envy.” What can their aim possibly be here? These observations not only do not challenge my characterization of leisure as nonpositional, they directly support it.

The key point is that if leisure is nonpositional, work becomes misleadingly attractive to individuals. Each person sees a chance to get ahead in relative terms by working a little harder (thereby, as Kashdan and Klein point out, seeming to be more important or more committed). Yet when all work harder, relative position remains unchanged—and hence the potential attraction of collective steps to limit work.

Surveys of associates in large law firms, for example, confirm that most respondents would prefer to work fewer hours at lower pay, but only if other associates also did so (Landers, Rebitzer, and Taylor 1996). Findings in the happiness literature also support the characterization of leisure as nonpositional. Even after controlling for other factors, higher
happiness levels are consistently associated with higher absolute levels of leisure (Kahneman et al., 2006).

The categories of behavior that society chooses to regulate reflect social judgments about the relevant deficits and excesses. Patterns of regulation are thus another source of evidence with respect to leisure’s nonpositionality. Almost all countries encourage leisure through regulation and social norms. Long before governments became involved, religions attempted to encourage leisure by designating Sabbath days on which work was forbidden. In the United States, the Fair Labor Standards Act encourages shorter working hours by its provision requiring premium pay for labor performed in excess of 8 hours per day, 40 hours per week, or on national holidays. European regulations are even stricter in their support of shorter hours. And many jurisdictions continue to enforce Blue laws, which make it unlawful for establishments to remain open during certain periods.

If so many countries actively intervene to constrain the number of hours that people would otherwise have chosen to work in unregulated markets, people must believe that working longer hours would reduce welfare. The conventional explanations offered for these regulations are completely unpersuasive. For instance, although France defended its recent requirement of a 35-hour workweek on the grounds that it was needed to stimulate jobs, such stimulus effects have never been demonstrated. Similarly, although many have defended hours regulations as needed to protect workers from employers with market power, the constraints imposed by such regulations bind most heavily for hourly workers in low-wage labor markets, which are among the most highly competitive by conventional yardsticks. In contrast, salaried workers in high-wage labor markets are relatively unconstrained by hours regulations, even though their employers are much more likely to occupy dominant market positions. The observed patterns of regulation are consistent with the hypothesis that leisure gets short shrift because of positional concerns. And contrary to what Kashdan and Klein suggest, support for such regulations does not signal any lack of commitment to excellence.

I once participated in a debate with a University of Chicago economist on the question of whether public policy should attempt to curb positional arms races. My opponent argued that it should not. In my opening presentation, I used the example of how perceptions of automobile performance are sensitive to context. Thus, for a car to be experienced subjectively as fast, it must accelerate more rapidly than most other cars in the same local environment. If people enjoy fast cars, the result is often an arms race focused on engine displacement. From the vantage point of
motorists collectively, this arms race is wasteful for the reasons discussed earlier. The taxes that many European countries impose on engine displacement might thus enhance welfare by inhibiting this arms race. And again, because we must tax something, only crybabies would think it relevant to object that taxation is coercive. Do Kasdan and Klein really believe that there are much better things to tax than engine displacement?

My Chicago opponent conceded that positional externalities might be important. But he argued that because we know so little about the specific ways they might matter, we are unable to predict even the direction, much less the size, of any distortions they might cause. To make his point, he asked the audience to imagine a community in which everyone valued having not a fast car but a slow one. To gain positional advantage, people in this community would then compete by buying automobiles with ever smaller engines. In equilibrium, cars would be “too slow” from the collective vantage point.

This argument is an example of what my former colleague Richard Thaler has called “sufficiency bias.” When economists afflicted with this bias confronts an argument they do not like, they rebut it by offering an argument that demonstrates the logical possibility of the opposite result. And having done so, they seem to consider the issue settled.

My Chicago opponent’s argument was logically impeccable but empirically absurd. Under his assumptions, the equilibrium automobile would indeed have been too slow. The only problem is that the kind of community he assumed does not exist. If we are looking for the right answers to important questions, that matters.

To their credit, Kasdan and Klein raise objections that are generally more thoughtful than those raised by my Chicago critic. But in fairness, I must also report that I detect a hint of sufficiency bias in many of them. Thus they claim, on the basis of no evidence other than some trend data on the amount of leisure in the US in recent decades, that we shouldn’t encourage leisure consumption because we don’t really know whether leisure is nonpositional.

Even in a state of imperfect information, however, it makes no sense to base policy on the assumption that we know nothing. The best available policies are those based on the best information we have. Positional externalities have been around for as long as animals have inhabited the planet. The historical record of human societies is replete with detailed information about them. Economists have been studying positional externalities formally for only a few decades. Even so, however, we already know quite a bit about them.
On conservative interpretations of the relevant evidence, the welfare losses from positional externalities are enormous, easily hundreds of billions of dollars a year or more. Savings of that magnitude could alleviate much human misery. Under the circumstances, it is a bold step indeed to counsel inaction on the basis of ignorance.

REFERENCES


ABOUT THE AUTHOR

Robert H. Frank is the Henrietta Johnson Louis Professor of Management and Professor of Economics at Cornell’s Johnson School of Management. His “Economic Scene” column appears monthly in The New York Times. His books, which include Choosing the Right Pond, Passions Within Reason, Microeconomics and Behavior, Principles of Economics (with Ben Bernanke), Luxury Fever, and What Price the Moral High Ground?, have been translated into ten languages. The Winner-Take-All Society, co-authored with Philip Cook, received a Critic’s Choice Award, was named a Notable Book of the Year by The New York Times, and was included in Business Week’s list of the ten best books of 1995. Falling Behind will be published in 2007. He is a co-recipient of the 2004 Leontief Prize for Advancing the Frontiers of Economic Thought. He was awarded the Johnson School’s Stephen Russell Distinguished teaching award in 2004 and its Apple Distinguished Teaching Award in 2005.
Work Incentives and Employment are the Wrong Explanation of Sweden’s Success

ANDREAS BERGH

CONTINUATION OF THE EXCHANGE BETWEEN ANDREAS BERGH AND PETER LINDERT FROM THE MAY 2006 ISSUE OF EJW.

Bergh Comment on Lindert (May 2006)
Lindert Reply (May 2006)

IN HIS RESPONSE TO MY COMMENT (BERGH 2006), LINDERT insinuates repeatedly that my criticisms of his book are little more than ideological bias. In this response, I will try even harder to recur to the facts. Still my conclusion is that Lindert is wrong about work incentives and employment in Sweden. To explain the so called free-lunch puzzle, we probably need to look closer at institutional quality and economic freedom.

I comment on Lindert’s remarks in order of their appearance.

Econometric evidence

Regarding the econometric evidence on the relation between government size and growth, Lindert writes “We should not push the econometric testing aside as if it were just a faulty camera that cannot take the picture Bergh knows to be true” (239). I didn't claim to know that the true picture is a negative correlation. In fact, I wrote:

* The Ratio Institute, Stockholm, Sweden SE-102 42.
Tax funded government activities include many things, some of which are probably bad for growth, whereas others are more likely to have a positive effect. Another possible explanation is that growth and big government are correlated because they are both at least partially the result of other omitted variables, such as certain socio-political traditions. (212)

On this point, I think that actual disagreement is very small.

Work-incentives for low-income earners

Lindert writes:

[Bergh] denies that Sweden has marginal rates as low as 30-35 percent (as I had said), arguing that benefit losses would give a job-getting single mother a much higher marginal tax rate. Given the complexity of quantifying all the marginal incentives, he could be right on this specific point. (242)

First of all, my comment made it very clear that the issue was work-incentives not just for single mothers (who are not necessarily poor), as Lindert suggests, but for poor and low-income earners in general.

Second, the complexity of the calculations is no excuse since the OECD has done most of the work. Their data confirms that work incentives, not only for single parent households, but for low-income earners in general in Sweden are much weaker than in the United States.¹

Lindert also says that the Sweden-US comparison should be made before the 1993 Earned Income Tax Credit reform. But, as one can learn from Lindert’s book, in 1967 the dollar-for-dollar benefit reduction of public assistance was abolished in the United States because it created poverty traps. It was reintroduced by the Reagan administration in 1981, and in 1993 the EITC reform once again strengthened work incentives substantially. Throughout this whole period, the Swedish social assistance has been reduced crown-for-crown, creating poverty traps that are still in place today. The upshot is that Sweden and the United States have

¹ OECD (2004).
sometimes had similar poverty traps, and most of the time since 1967 work incentives have been stronger in the United States.

For some convincing empirical evidence that the structure of economic work incentives for low-income earners created by different benefit reduction schemes have substantial behavioural effects, see Munts (1970).

In his reply, Lindert notes that the provision of health care, child care, and other services can be included when comparing work incentives. This would strengthen his case, because in Sweden these services are less means-tested against income than in the United States (242). However, first, while these services are highly subsidized in Sweden, they are not free. For example, child care and dental care are not negligible expenditure for low-income earners. But more importantly, if needed, these costs are covered by social assistance, adding to the poverty trap. Thus, including these services when comparing work incentives for poor in Sweden and the United States strengthens my initial point that Lindert has missed the poverty trap created by social assistance in Sweden. Lindert should acknowledge that he overlooked something important when he overlooked Sweden’s social assistance program.

**Work incentives for high-income earners**

Lindert goes on:

While there lingers some uncertainty about those low-income single mothers, [Bergh] surely owes us an explanation of why he passed up all the evidence that top income earners [in Sweden] did not pay much higher rates than middle earners, either before or after the tax reform of 1991. Before the reform, they had huge opportunities to pay less than the high top rates that got into the media, and the reform slashed those top rates. Nor does he mention the important point that Sweden levies low corporate income taxes. (243)

Lindert's argument here is that before the 1991 reform high-income earners could get away with not paying the marginal tax rates. I agree, but this does not change the fact that statutory tax rates for high-income earners were very high. Furthermore, the mechanism Lindert points to
means that high-income earners spend resources on tax-avoidance, resources that could have been used more efficiently. This is one of the costs associated with high and progressive taxes.

After the tax-reform, marginal tax rates for high-income earners are lower, but they are still higher than for middle-income earners. Furthermore, as I discuss in Bergh (2004), high-income earners in Sweden not only pay an extra 20 or 25 percent in state income tax, their incomes are also not fully covered by the social insurance systems, because there is an upper limit on the positively income-related benefits. For this reason, the disposable income curve, as shown in Figure 1, has three segments: The poverty trap created by social assistance, a segment with moderate work incentives with marginal effective tax rates below 50 percent, and a section with lower work incentives where the marginal effective tax rate is above 50 percent. The difference created by the state income tax is always 20 percentage units or more (as can be seen in Figure 2, page 216 in my first comment). The structure of marginal taxes can easily be confirmed by OECD (2004).

When Lindert moves on to note that taxation of corporations is low in Sweden he is right, but he is no longer talking about work incentives for employed workers. Furthermore, the combination of high labor taxation and lower taxation of corporations produces incentives to keep profits within the corporation rather than to pay wages—an effect that is not obviously desirable.

Employment of women

I remarked that Lindert’s picture of high gender equality in Sweden is biased because he does not take into account that Swedish women often work part-time. Lindert responds by asking about the value of household work:

Are we to imagine that women are forced into useless partial idleness because of taxes? Do they knock off formal work at 3:00 pm because taxes are too high for a full day, or quit their jobs for part of the year and get a new job later to cut taxes? Instead, should we not look at the greater part-timing and partial unpaid work for women in Sweden (and Netherlands and elsewhere) as a response to a better opportunity set? (244)
Lindert may well choose to view greater part-timing and unpaid work as a response to a better opportunity set. My point was much simpler: Displaying only employment ratios is not a convincing way to explain why the welfare state seems to have no negative effect on GDP. Household work is not included in GDP. My noting this says nothing about my personal opinions regarding the value of unpaid work and idleness.

Lindert’s rhetorical questions make fun of my argument, as if Lindert did not understand that my point was the following: High labor taxation means that it is expensive for a person to hire another person to do a job. One reason why Swedish women work part-time is that there is no functioning market for household services. One major reason is that high
labor taxation creates a do-it-yourself incentive—not only in theory but also in reality, as verified for example by Davis and Henrekson (2004).

Lindert wants me to explain why I am skeptical about his explanation of high female employment in Sweden based on high female labor supply elasticity (244, footnote 12). I am skeptical because, while labor supply elasticities describe how the number of hours worked depends on the net wage, Lindert does not use the number of hours worked but employment ratios. I clearly stated (220) that when we consider the number of hours worked instead of employment ratios, female employment in Sweden is actually lower, not higher, than in the United States. To explain the two facts that Sweden has a higher female employment ratio, but lower number of hours worked, Lindert’s elasticity story does not work.

If Lindert’s elasticity story were to work, presumably it should work in both the public and the private sectors. To help explain my doubts about Lindert’s story, I noted that the female employment expansion in Sweden has been entirely in the public sector. Rather than address this matter, Lindert simply acknowledges the fact (which was strangely omitted from his book), and diverts us with another ideology play: "Are readers supposed to shake their heads in knowing disapproval, because they all know that public employment is inferior to private?"

Retirement

Lindert writes, "for women, the share [in Sweden] still working in the 55-64 age bracket is also high by international standards, despite Bergh's claims to the contrary" (244-5).

I did not claim the contrary. I wrote that "[Lindert] has not verified that the high employment ratio in the age group 55-64 implies that Swedish elderly actually work more hours" and that Lindert is wrong to claim that "Swedish women work to later average retirement ages than women anywhere else in the world" (Lindert 2004, 289), noting as I did that in fact women in Iceland, Mexico, Korea, Ireland, Japan, Portugal, Switzerland, United States, Norway, and Denmark work to a later age (227). Again, I simply recurred to the facts, and Lindert never answered the challenge.
Economic freedom

In my reply, I suggested that a number of institutional reforms may explain the perhaps surprisingly good economic performance of Sweden and other high tax societies. Lindert comments:

Most of the article tries to stress that Sweden has painted itself into a corner today, yet this part emphasizes that Sweden reformed its way out of problems: [...] It is hard to make out what this final section is trying to achieve, beyond putting in a plug for subjective indices of "freedom" by the Heritage Foundation and the Fraser Institute. (245-246)

First of all, the index developed for the Fraser institute by James Gwartney and Robert Lawson has been used for serious research published in many established scholarly journals, as well as in work by the IMF and the World Bank. Indices are never perfect, but many scholars seem to agree that this is a good attempt to quantify various aspects of institutional quality and economic freedom.

Second, Lindert puts freedom in sneer quotes, without explanation. I find this curious. I wonder how Lindert reads Adam Smith's economic treatise on the meaning, worthiness, and limitations of natural liberty. It is undisputed that, as Smith argued, economic freedom has a big and robust positive effect on growth. Doucouliagos and Ulubasoglu (2006) conduct a meta-study of 52 studies dealing with the impact of economic freedom on economic growth, and conclude that “economic freedom has a robust positive effect on economic growth regardless of how it is measured” (68).

In Bergh (2006) I develop the idea that big welfare states can use economic freedom to compensate for negative effects caused by big government and high taxes. Figure 2 illustrates Sweden’s development by showing the economic freedom index average for legal structure and property rights, access to sound money, freedom to trade and regulations of credit, labor and business as well as total taxes relative to GDP. This collection of components is the total index except for the government-size component.
I think it is clear that reforms toward increasing economic freedom could well be one of the keys to the so-called free-lunch puzzle. It’s simple enough, but Lindert says he doesn’t see the point.

Concluding remarks

In the final section, Lindert goes on to claim that he agrees with me on the importance of reforms in Sweden. He goes on to remark about the negative effects in Europe of employment protection laws, product market regulation, and the centralization and under-pricing of Sweden’s university education (247).

Lindert mentions all this because he wants to persuade me and the readers that “[t]he welfare state is the wrong target” (246). To some extent, I agree with him on this point. Throughout his reply, Lindert presumes that I am an opponent of the welfare state. Since Lindert predicates so much of his reply on my presumed ideological sensibilities, it seems appropriate that I disclose what they actually are. I am in fact rather friendly to the welfare
state, and, otherwise, to economic freedom. In my view, the Swedish welfare state creates problems in employment and work-incentives, problems that are much bigger than they have to be. I would never claim that the welfare state is a free lunch.

REFERENCES


Go To Second Reply by Peter Lindert
Second Reply to Bergh

PETER H. LINDERT*

I AM PLEASED THAT IN HIS SECOND COMMENT ANDREAS BERGH (2006b) has decided to “try even harder to recur to the facts,” and agrees that on the key econometric front our “actual disagreement is very small” (pp. 452, 453). He also now reveals that he is “in fact rather friendly to the welfare state.” We seem to have achieved some convergence.

Still, he retains some lingering disagreements with my analysis, and two main disagreements deserve to be noted briefly. The common element in these disagreements is the basic clash that I noted both in my book (Lindert 2004) and in my first reply to Bergh (Lindert 2006): While they concede that the econometrics and the raw correlations fail to show any negative GDP effect or growth effect of higher total social transfers, some economists, like Bergh, insist that the costs predicted by their theories must be there.

TAXES AND FREEDOM

“I think it is clear that reforms toward increasing economic freedom could well be one of the keys to the so-called free-lunch puzzle. It’s simple enough, but Lindert says he doesn’t see the point” (Bergh 2006b, 459). Readers wanting to see the point should look again at his new chart on taxes and the freedom index. What Bergh apparently wants to show is that the rise of economic freedom since 1970 was bigger for Sweden than for the USA or France, with the United States retaining a lead as of 2000. What stands out in the diagram is that Sweden’s much higher tax ratio is consistent with

* University of California, Davis. Davis, CA 95616.
economic freedoms that are both rising and high by the standards of the USA of 1970 and earlier. So if Bergh feels that “reforms towards increasing economic freedom could well be one of the keys to the so-called free-lunch puzzle” (p. 459), he should perhaps point out that the high and rising tax shares that concern him have not dragged down economic freedom or Sweden’s success.

The coexistence of high tax rates and a high freedom index in Sweden is all the more remarkable since the freedom index Bergh uses (Gwartney and Lawson 2005) was designed to give lower freedom marks for countries with high taxes and big government. The leading component of the Fraser Institute “economic freedom” index loads on the following factors (9):

1: Size of Government: Expenditures, Taxes, and Enterprises
   A. General government consumption spending as a percentage of total consumption.
   B. Transfers and subsidies as a percentage of GDP.
   C. Government enterprises and investment as a percentage of GDP.
   D. Top marginal tax rate (and income threshold to which it applies).
      i. Top marginal income tax rate (and income threshold at which it applies)
      ii. Top marginal income and payroll tax rate (and income threshold at which it applies)¹

Thus dragged down in the index by design, Sweden nonetheless has ranked close to the United States in economic freedom ever since the index was first calculated for 1970. Alternatively, one could have used an alternative index of good government institutions, one that doesn’t define bigger government budgets as a loss of freedom. One such index is the clean government rating given by Transparency International’s Corruption Perceptions Index, which found that all four Nordic countries were consistently at the top of the clean-government ranks (with Singapore and New Zealand, as of 2003), while the United States is tied for 18th and 19th (Transparency International 2003). Whichever institutional rating one tries, Sweden’s high tax take has

¹ Note that these top marginal tax rates featured in the freedom index are the same ones that Bergh agrees were seldom paid by Swedes before they were lowered in 1991. Thus part of what passes for a “freedom reform” that Bergh sees in his Swedish freedom graph may simply reflect Sweden’s removing these ineffectual high rates.
not been inconsistent with good institutions, either before or after the "reforms toward increasing economic freedom" (Bergh 2006b, 459).

Given that freedom and clean government seem so consistent with the welfare state and high taxes, one might wonder how it is that Doucouliagos and Ulubasoglu (2006) concluded, as Bergh cites them, that "economic freedom has a robust positive effect on economic growth regardless of how it is measured" (458). They got this plausible result from global samples, not from samples of OECD rich democracies. The positive effect of economic freedom on economic growth comes from the bad performance of dysfunctional dictatorships like those in Zaire and Zambia. The same inclusion of the world’s worst in global samples also explains how some economists find a negative effect of government consumption (not social transfers) on GDP growth. We could call it the Mobutu effect. It says nothing about the effect of social transfers or the welfare state in a context of highly educated democracies. Adam Smith is right about freedom, but freedom can’t explain much here simply because it doesn’t differ much between OECD countries.

On the alleged effect of Sweden’s high statutory top tax rates before 1991, Bergh again lapses back into theory when the facts fail to produce for him. Having agreed with me that “before the 1991 reform, high-income earners could get away with not paying the marginal tax rates” he still feels that Sweden’s tax system must be costly: “but this does not change the fact that statutory tax rates for high-income earners were very high. Furthermore, the mechanism Lindert points to means that high-income earners spend resources on tax-avoidance, resources that could have been used more efficiently. This is one of the costs associated with high and progressive taxes” (455). We are back to that same I-see-a-tax-wedge reasoning that I questioned both in the book and in my earlier reply to him. It is one thing to have a tax rate on the books and another to show that people respond to it. He shows no statistical evidence that a significant share of GDP was lost in tax-avoidance activities.

POVERTY TRAP

Bergh’s second comment steps up the frequency of his references to a “poverty trap created by social assistance in Sweden” (454). That will seem odd to anybody who knows the poverty rates measured by the OECD and by the Luxembourg Income Study. Using comparable measures
of absolute purchasing power parity for the year 2000, one finds that only 5.3 percent of the population was below the poverty line in Sweden, versus 17.2 percent falling below the same absolute poverty line in the United States. How did Sweden achieve that? It might have something to do with Sweden’s willingness to redistribute to the poor. Figure 1 shows the well-known correlation between countries’ efforts to help the poor with social spending and their relative success in cutting poverty.

What Bergh seems to mean by the “poverty trap” is simply that statutory rates of taxation and withdrawal of benefits imply that the poor face an ostensibly strong incentive to avoid work. Once again, he reverts to that I-see-a-tax-wedge reasoning. We need evidence that people respond to it. Econometric studies have consistently shown that the employment response to ostensibly large tax wedges is very small in today’s democracies.
The alleged “evidence” of large work effects comes overwhelmingly from theory and not from any robust econometrics, despite his new claim that Ray Munts’s 1970 study showed big effects.

The welfare state remains a “free lunch” in the same sense that I used this phrase in my 2003 NBER paper, my 2004 book, and my first reply to Bergh earlier this year: The higher social transfers of the welfare state have brought less poverty, less inequality, and longer life expectancy with no statistically significant cost in terms of GDP.

REFERENCES


Bergh, Andreas. 2006b. Work Incentives and Employment Is the Wrong Explanation of Sweden’s Success. Econ Journal Watch 3(3): 452-460. Link


A Little More Liberty: 
What the JEL Omits in Its Account of 
What the Economic Report of the 
President Omits 

DANIEL B. KLEIN AND MICHAEL J. CLARK* 


Abstract

It is the highest impertinence and presumption, therefore, in kings and ministers, to pretend to watch over the economy of private people. . . . Let them look well after their own expense, and they may safely trust private people with theirs. 

—Adam Smith (1776, 346) 

ADAM SMITH HELD THAT THE POLITICAL PROCESS—INVOLVING the mentalities of ordinary citizens, ministers, politicians, and intellectuals—is prone to an under-appreciation of the relative virtues of natural liberty and the social order that it generates. He attacked many economic policies by explaining how free individuals advance the interests of society. By describing how the independent actions of many individuals would generate 

* Department of Economics, George Mason University. Fairfax, VA 22030. 
For assistance with developing the survey of George Mason University professors, we thank Ted Balaker, Carl Close, Jerome Ellig, Adrian Moore, and Jane Shaw. For other suggestions we thank Erin Bartee.
beneficial order, and how government is riddled with knowledge and incentives problems, including vanity, conceit, and self-delusion, Smith centered political economy on the presumption of natural liberty. (That presumption also prevailed among French economists from the time of the Physiocrats.)

In his famous JPE article “Adam Smith and Laissez Faire,” Jacob Viner explored Smith’s apparent inconsistencies and exceptions to natural liberty, but concluded authoritatively: “There is no possible room for doubt, however, that Smith in general believed that there was, to say the least, a strong presumption against government activity beyond its fundamental duties of protection against its foreign foes and maintenance of justice” (Viner 1927, 219). The presumption of liberty suggests that interlocutors ought to regard the liberty maxim as the normal and proper guide for public policy. The liberty maxim speaks plainly and directly to many important issues. The Smithian presumption places the burden of proof on those siding with contraventions of that maxim.

In September of 2005 the Journal of Economic Literature published a review of that year’s Economic Report of the President. Five authors, writing separate sections of the review, raised numerous omissions in The Economic Report of the President (ERP). They had an opportunity to object to current government interventions, an opportunity to challenge the “impertinence and presumptions” of “kings and ministers [who] pretend to watch over the economy of private people.” The ERP itself tepidly proposes a few liberalizations, and regarding those few the JEL authors are sometimes supportive and sometimes dubious. But what is most significant is that the JEL authors do not suggest a single liberalization (outside the few raised by the ERP).

There is little trace of the Smithian character. The failing is especially noteworthy because an image of Adam Smith graces the cover of the Journal of Economic Literature.

**SOME CONCRETE EXAMPLES OF POSSIBLE LIBERALIZATIONS**

First, we wish to put some concrete possible liberalizations onto the table. We conducted a modest little survey to generate a ranking of liberalizations. We sought a convenient sample of economists who could be
taken to be rather Smithian in their sensibilities, and turned to the faculty of our own department at George Mason University. We asked them to rank 57 possible federal liberalizations in terms of how worthwhile each would be as a reform proposal in the ERP. Appendix 1 provides a full account.

The reform proposals were divided into 35 deregulations and 22 privatizations. Of the 35 deregulations, the top ten were: diminish trade restrictions, reduce agriculture subsidies and regulations, reduce FDA restrictions, reduce anti-trust enforcement and restrictions, reduce regulations on healthcare facilities and professionals, repeal restrictions on competitive mail delivery, liberalize drug prohibition, repeal laws that require banks to keep tabs on customers and report activity to the government, revisit Sarbanes-Oxley, and liberalize anti-discrimination laws. Of the 22 privatizations, the top ten were: disaster insurance, the U.S. Postal Service, Amtrak, Social Security accounts, space exploration, power and electric infrastructure, job training and workforce assistance programs, disaster response agencies, water infrastructure, and federal loan programs. The complete rankings (and exact wording) are found in Appendix 1. The rankings by themselves do not deserve analysis or commentary. The 57 liberalizations and the coarse rankings provide a handy list of possibilities when examining the ERP and JEL review.

THE 2005 ECONOMIC REPORT OF THE PRESIDENT

Each year the President’s Council of Economic Advisors prepares The Economic Report of the President (the ERP). It offers an overview of the condition of the economy and purports to discuss policy issues that are important and timely. Being a government report prepared by a partisan executive office, the ERP is one of the last places one would expect to find serious challenges to the governmental status quo. It would be natural to expect that the report’s authors would avoid liberty and kindred ideas and instead work from a presumption of the status quo. The status quo exists, and officials need to presume that it does so for good reasons. Rejection of

---

1 For several reasons: (1) We don’t know how respondents weighted the various aspects (e.g., potential efficiency gains, amenability to economics, being appropriate to the ERP); (2) the 1, 2, or 3 ratings only coarsely capture judgment; (3) although most of the GMU economists are quite liberal in the original sense, it cannot be assumed that all the respondents are; (4) eleven respondents is less than 50 percent of the survey population.
the status quo would require debate and justification, as well as the admission that the existing policy was a mistake, that the political process went wrong.

Yet the 2005 ERP, which contains a “Letter of Transmittal” signed by the three members of the Council of Economics Advisors (Chairman N. Gregory Mankiw, Kristen J. Forbes, and Harvey S. Rosen), pays a good deal of attention to the benefits of a free economy. The ERP projects the view that the government, with a focus on protecting property rights, could promote the nation’s wealth by allowing individuals to handle their own affairs. In chapter 6, the ERP takes a general free-market stance:

An innovation will succeed if it passes the market test by profitably delivering greater value to customers. Successful innovations blossom, attracting capital and diffusing rapidly through the market, while unsuccessful innovations can wither just as quickly. In this way, markets allow capital to flow to its highest-valued uses. This engine of growth can falter, however, if government policies distort the market signals that guide innovative activity. (ERP 135)

In a number of chapters, the ERP projects a free-market message and offers some standard argumentation for basic free-market views.

Even though the ERP’s arguments are often couched in generalizations about the gains from property rights, the reader senses the constant hesitation to take on particular agencies or policies. If private property and free markets work so well, it would seem that the authors should be challenging many federal policies.

The list of 57 federal liberalizations was purposely made to represent the policy views of an economist whose judgments favored private property and freedom of contract. If the Smithian character were the uninhibited mindset of the ERP, one would expect it to treat a fair number of the 57 liberalizations.

We have detailed ERP’s page-by-page discussion of issues and its judgments. The analysis is contained in an Excel file linked from Appendix 2.

The ERP directly treats and favors four of the 57 federal liberalizations. First, an entire chapter treats international trade, and the general argument is for free trade. The chapter focused directly on trade agreements that could expand markets worldwide, but there is no mention of unilateral removal of any current U.S. tariff or import quota. Second, an entire chapter treats immigration and, again, projects a general call for liberalization. The chapter
KLEIN AND CLARK

reports that “the benefits of immigration exceed the costs” (*ERP* 93). But again the chapter does not even hint at a general call for relaxing immigration controls; rather it delves only into the President’s temporary worker program. Third, in the chapter entitled “Expanding Individual Choice and Control,” which focuses on the virtues of property rights, the *ERP* discusses the eventual failure of the current Social Security system and the gains from ownership and control of retirement funds; it advocates creating personal retirement accounts. Fourth, in the same chapter, the *ERP* calls for using a system of individual transferable quotas (ITQs) to regulate fisheries.

As for other reforms on our list of 57 liberalizations, the *ERP* vaguely approached three others. It discussed the gains from a flexible labor force, but didn’t specifically mention current union privileges or the minimum wage. Regarding telecommunications, the *ERP* argued for a strong enforcement of spectrum rights and regulatory adaptation to changing market conditions. Finally, the *ERP* spoke very highly of tradable pollution permits, but did not suggest any general relaxation or reform in environmental controls.

Our survey list of 57 liberalizations was drawn up in light of what the *ERP* contained. We included any liberalization mentioned in the *ERP*, but with three quasi-excceptions. In the “Expanding Individual Choice and Control” chapter, the *ERP* offers as property-rights applications health savings accounts and “Millennium Challenge Accounts” for foreign aid (we excluded these reforms because of their murkiness qua liberalization). More significantly, the chapter gives a lengthy and very positive account of school vouchers—perhaps the authors found it easier to discuss radical liberalization when it was not a federal-government issue (the reason we excluded it). The *ERP* included no other reforms that could possibly be viewed as liberalization.

Again, we should expect administration economists to produce a tepid report centered on the status quo, remaining between the 45-yard lines, and devoid of criticism. What we should expect in a review of the report by academic economists is another matter.

---

2 The section can be found on pages 127-129.

3 Although the idea of displacing Social Security with personal retirement accounts would seem to be liberty enhancing, political dynamics make the creation of personal retirement accounts a hazardous step, as the displacement is never ensured and as it would likely deepen government involvement and regimentation of the private-investment market.
THE \textit{JEL} REVIEW OF THE \textit{ERP}

As a preface to the \textit{JEL} review, the \textit{JEL} Associate Editor James Levinsohn provides an editorial note which explains that the review consists of five separately authored parts:

\begin{itemize}
  \item \textbf{Joel Slemrod} [University of Michigan] reviews the discussion of tax reform.
  \item \textbf{Joe Farrell} [University of California, Berkeley] reviews the ERP’s chapter titled: ‘Innovation and the Information Economy.’
  \item \textbf{Gordon Hanson} [Massachusetts Institute of Technology] reviews the chapters on international trade and on immigration.
  \item \textbf{Robert Hall} [Stanford University] reviews the discussion of the adverse macroeconomic impact of rising oil prices.
  \item \textbf{Jonathan Gruber} [University of California, San Diego] reviews the ERP chapter titled ‘Expanding Individual Choice and Control.’
\end{itemize}

(Levinsohn’s editorial note to Farrell 2005, 801; bold added)

Collectively, the several authors treat most of the \textit{ERP}, especially its most policy relevant chapters. Thus, they would be in a natural position to point out that the \textit{ERP}, while touting property rights and free markets, fails to apply those principles to the many deregulation and privatization ideas deemed important by Smithian economists. By challenging the \textit{ERP}, the authors could have helped to provide a public check on government activity. The idea of an economist providing a public check on government leads back to Smith’s presumption of liberty. That presumption set the stage for economists to promote good economic policy by countering the conceit and folly of the politically powerful, the intellectuals, and the lay public alike.

The \textit{JEL} authors criticize the \textit{ERP} mostly for its omissions. But the bottom line is that they have no criticism whatsoever for the \textit{ERP}’s neglect of possible liberalizations.

We have itemized the critical points raised by the \textit{JEL} authors, including what they see as the \textit{ERP}’s errors of omission. The itemization is linked from Appendix 3. The list of omissions shows what the \textit{JEL} authors felt the \textit{Economic Report of the President} was leaving out. It is the \textit{JEL} authors’ check on the government’s oversights. Appendix 3 contains all the details, but the main point is what the authors omit, so here we remark only very briefly on what each says.
Robert Hall’s section deals solely with the ERP’s claims about the macro effects of increased oil prices. Those claims play a minor role in the ERP. Hall discusses macro theory and says nothing about policy. A senior fellow at the conservative Hoover Institution, Hall coauthored a well-known book espousing the flat tax and served President Ronald Reagan. One can only guess why his remarks about the ERP are so narrow.

Joel Slemrod reviews the ERP chapter “Options for Tax Reform.” Since our concern here is potential liberalizations in the form of deregulation and privatization, tax reform is not germane. Slemrod accepts the ERP’s estimate that the excess burden of raising another dollar in income tax is somewhere between 30 and 50 cents, not counting compliance costs, which he suggests “may be as high as 10 percent of tax revenues” (817). He focuses on the ERP’s ideas about simplifying the tax code and moving to some form of consumption tax. Slemrod says (reasonably, to our ears) that the chapter emphasizes the evils of taxation yet neglects how to negotiate the trade-offs so as to fulfill the mandate of tax-revenue neutrality. Also, he doubts that tax cuts succeed in “starving the beast” and takes the ERP to task for not linking to deficits and the necessity of future taxes. Slemrod takes up the ERP’s discussion of shifting to consumption taxes. The points are detailed in Appendix 3. To our inexpert ears, Slemrod’s commentary seems insightful, expert, and certainly not unreasonable. Our criticism in this article is not directed at him.4

Jonathan Gruber reviewed the crucial ERP chapter titled “Expanding Individual Choice and Control,” which emphasizes property rights as an approach to policy reform. Gruber supports the section on tradable pollution rights, but otherwise says the property-rights approach was oversold. He sees problems with applying property rights to insurance and services markets:

First, services markets are much more subject to informational failures, giving birth to the twin demons of adverse selection and moral hazard, market failures that property rights can do little to address. Second, service markets are typically marked by a higher share of costs in administration/servicing, and these costs are minimized by the large pools that can be provided by government programs. (806)

4 If we were to get into omissions on the front of desirable tax reform, we would perhaps be inclined to raise the idea of shifting very gradually to “geo-rent” taxation, that is, taxation of what land devoid of improvements would rent for; see Foldvary 2005.
Gruber casts doubt on the ERP's endorsement of school vouchers, concluding: “the jury is still out on the systemwide effect of vouchers” (807). Next he casts doubt on applying property rights to Social Security and health savings accounts, raising adverse selection. For details, again we refer the reader to Appendix 3. In general, Gruber says the ERP does not sufficiently highlight the market failures brought on by property rights and privatization.

Gordon Hanson reviewed the ERP chapters on immigration and international trade. Although Hanson never judges against freer trade and immigration, he raises a number of criticisms, several of which tend to counter the ERP's pro-immigration and pro-free trade stance. He says that the ERP: (1) oversold the case for international trade by overemphasizing the benefits from inward foreign direct investment (813), (2) did not adequately discuss the distributional consequences from globalization (813-814, 816), (3) misrepresented flexibility in the market for immigrant labor (814), (4) perhaps under-estimated immigration's adverse effect on native low-skilled wages (814-815), and (5) misrepresented the fiscal impact of immigrants (815). Overall, Hanson felt the ERP's discussion on immigration and trade put forth an overly optimistic view of globalization.

Joseph Farrell reviewed the ERP chapter “Innovation and the Information Economy.” He makes several critical points, detailed in Appendix 3, and here we mention just a few. He says the ERP is dismissive and remiss regarding the benefits of regulation and non-property rights alternatives, for example, neglecting that the “Internet and e-mail were developed by the government and universities on a noncommercial basis” and that the “world-wide web was developed at a European government research institute” (802). Farrell makes some comments which make us wonder about his understanding of the logic of property ownership and contract. He suggests that open-source software demonstrates that the “strong-property-rights model” is not always appropriate or desirable (802). Well, perhaps, but the point of minimally attenuated private ownership of property and the freedom of contract (Smith’s natural liberty) is that they offer a foundation for spontaneous order and an alternative to government intervention, and clearly open-source proceeds under such principles. More troubling is Farrell’s assertion that “people have a property right to the security of their data” and suggestion that the credit rating agency be made liable “for full compensatory damages when it skimps on employee screening, treats identity thieves as customers, or laxly keeps my social security number unencrypted” (805)—a suggestion he fashions a property-rights reform. We reject the notion of individual’s having ownership of “the security of their data,” seeing such issues rather as matters of contract, and in fact count...
Farrell's proposal as interventionist, as it would entail incursions on the free nexus of private property and contract.

But what really matters about the *JEL* review is what is not found. The authors do not suggest as worthy reform a single ERP-omitted potential liberalization. As the ERP mentions only a few liberalizations, this zero is striking.

**WHAT WE AS ECONOMISTS KNOW**

In his editorial note, James Levinsohn elaborates on the instructions given to the reviewers:

The *ERP* in principle should provide an accurate assessment of the consensus professional views of economists on any given issue, based on the research to date. Does the discussion in the *ERP* in fact accurately summarize what we as economists know? (Levinsohn editorial note to Farrell et al 2005, 801)

Levinsohn invokes an idea of scientific consensus that we think is specious. Political economy differs from more established sciences in many ways, particularly in the ways that the purposes of the discipline are determined, pursued, funded, realized, and validated. In the physical sciences there will be some issues like global warming that are politicized, but in political economy the important issues are necessarily politicized, not only because the topics involve policy, but because the reform process (including the persuasion process) is a matter of politics (whereas treating a disease or dating a solar system generally is not). Levinsohn should realize that political economy is bound up with deep-seated ideological sensibilities, notably when it comes to judging what the most important things are and what the discipline has to say about them. In a politicized discipline, it is entirely natural that “what we know” is heterogeneous and conflicted, and that the majority of putative scientists might be wrong on important things.

The ideological nature of political economy is revealed by an empirical result that is now pretty well established. On many important policy issues, such as the FDA, there is a significant impasse between what most economists think and what economists who actually write on the issue and express a policy judgment think. Surveys of economists (e.g., Whaples...
2006; Fuller and Geide-Stevenson 2003) give us a general reading of what economists in general think on specific policy issues. Going back to Kearl et al. (1979) and extending to other countries such as Ricketts and Shoesmith (1990; 1992), the various surveys generally paint the same picture of policy opinions among economists in general. On the other hand, special investigations of policy judgments by economists who actually write on and judge the issue often come to a noticeably different conclusion. On the FDA, for example, while most AEA members are generally supportive of current restrictions (Klein and Stern 2005; 2006), economists who research and publish policy judgments mostly favor liberalization, often to a great degree (Klein and Tabarrok 2006). (The George Mason economists ranked FDA liberalization third of 35 deregulations.) This journal (Econ Journal Watch) contains a section entitled “Do Economists Reach a Conclusion?” to determine whether the economists writing on and publicly judging an issue point in a specific reform direction. Many of the articles show that among such expert economists there is significant consensus in favor of liberalization. However, on many issues, notably the FDA, so-called consumer protection, and drug prohibition, the average AEA member seems to be significantly more aligned with the status quo.

So, when Levinsohn asks whether the ERP accurately summarizes “what we as economists know,” how do we identify what we as economists know? If the consensus of economists in general disagrees with the consensus of economists who study the FDA etc., then which consensus represents what we as economists know? If the latter, then we should ask why it is that on many issues most economists do not know “what we as economists know.”

WHY DID THE JEL OMIT LIBERALIZATION?

The JEL’s failure to suggest a single ERP-omitted liberalization brings us to the most interesting question of all: Why?

Here, one must bravely face issues of ideological heterogeneity. A variety of research—based on surveys, campaign contributions, and voter registration—establishes that the ideological dispositions of most AEA members are like those of most social science and humanities professors, namely, socially democrat, though somewhat less so. If we may suppose that the people principally responsible for the JEL review are of that
character, a supposition supported by McEachern (2006) and Klein (2006), then it becomes easier to understand the neglect of liberalization.

In our judgment, a great example of terrible federal policy is drug prohibition, which has disastrous consequences falling disproportionately on people who are least well off. The George Mason economists ranked drug liberalization seventh of 35 deregulations, and Thornton (2004) shows that most economists who write about the issue favor liberalization. As regards “Expanding Individual Choice and Control,” a good place to start would be repealing prohibition and releasing the hundreds of thousands caged for consensual crimes. Yet the JEL authors make no mention of prohibition. A primary reason social democratic intellectuals do not stand up against such policies is that such policies are the status quo. To condemn them would be to admit that American democracy can, for generations, go obviously and terribly wrong. Social democrats affirm politics and government as forms of collective experience and validation, and it would be very awkward to affirm a process while admitting that it persistently makes stupendous errors. Once the social democrats admit one gross error, they open themselves up to many embarrassing questions, including the following:

- What were the mechanisms that drove such error (and lack of correction)?
- Have those mechanisms operated generally?
- Have they routinely produced and preserved policy disasters?
- Do we live in a polity riddled with unenlightened policies?
- Does the livelihood of yourself or friends or family depend on unenlightened policy?
- Does support for social democracy lend horsepower to the pathological mechanisms?

It is our sense that one reason social democrats and other establishment intellectuals hew so closely to the status quo is that they need to keep all such questions taboo. One way to do so is to ignore liberalization proposals outside the 45 yard lines.
CONCLUDING REMARK

The *JEL* authors missed a chance to challenge bad status-quo policies. Ronald Coase wrote: “To ignore the government’s poor performance of its present duties when deciding on whether it should or should not take on new duties is obviously wrong.” The result is “an ever-expanding role for the government in economic affairs” (Coase 1999, 50, 51). Excessive encroachment on natural liberty was exactly the problem that the original liberals sought to combat.

In his editorial note, James Levinsohn announces that “beginning this year, the *JEL* will be reviewing the *ERP*” (801). We look forward to the future reviews of the *ERP*, and hope that they will begin to do justice to the image stamped on the cover of the journal.

APPENDIX 1: Survey about ranking federal liberalizations in importance, administered to faculty members of the George Mason University department of economics.

*Downloads:* 1) the survey (pdf); 2) the survey results based on the 11 responses (Excel).

DISCUSSION OF THE SURVEY

We gave faculty members of the George Mason University department of economics the survey containing the possible federal liberalizations. We asked them to rank them in importance by rating each issue 1, 2, or 3. The survey directs the respondent to imagine that he or she has been hired by the Council of Economic Advisors to work as a principle writer of the *ERP*: “You are basically free to write and publish in the Report what you want. Please rate which issues you feel would be most important to discuss, given the purpose and context of *The Economic Report of the President*.” The survey elaborates by telling the respondent to keep three aspects in mind: (1) The issue’s importance, in terms of potential efficiency gains; (2) The
value of creating public awareness of the issue; (3) The applicability of economic principles to the issues (specifically principles that usually recommend liberalization/privatization). We did not coach the respondents in any way or say anything about the JEL review.\textsuperscript{5} The reform proposals were divided into a set of 35 liberalizations of regulations and a set of 22 privatizations. In creating the lists, we consulted various guides to federal policy and collected reform ideas that seemed important and potentially salient as applications of economic principles. The lists are far from complete. Sufficiency calls simply for a wide smattering of issues. We focused on deregulations and privatizations. We avoided redistribution programs, tax-reform issues, and military policy.\textsuperscript{6} For each section, the respondent was to mark each reform 1, 2, or 3, with 1 for most important, and to distribute the three ratings in roughly equal proportions. The survey was completed by 11 GMU economics professors. Of the 35 liberalizations, the ranking emerging from the responses is shown in Table 1.

\begin{table}[h]
\centering
\caption{The Federal Liberalizations Most Worth Discussing in the ERP, According to 11 GMU Economists}
\begin{tabular}{|c|l|}
\hline
\textbf{Rank} & \textbf{Liberalization of Federal Government Regulation} \\
\hline
1\textsuperscript{st} & Diminish trade restrictions (tariffs and quotas). \\
\hline
2\textsuperscript{nd} & Phase down all agricultural subsidies and liberalize agricultural regulations. \\
\hline
3\textsuperscript{rd} & Reduce FDA regulations on pharmaceuticals, devices, and information. \\
\hline
4\textsuperscript{th} & Lighten the Federal Trade Commission and Department of Justice’s anti-trust enforcement and restrictions. \\
\hline
5\textsuperscript{th} & Reduce regulations on healthcare facilities and professionals. \\
\hline
6\textsuperscript{th} & Repeal legal restrictions on competitive delivery of mail. \\
\hline
7\textsuperscript{th} & Decrease the size of the Drug Enforcement Administration in conjunction with liberalizing the drug laws. \\
\hline
\end{tabular}
\end{table}

\textsuperscript{5} Klein did not participate in the survey.

\textsuperscript{6} This was done partly to simplify the matter and partly because the core insights of economics—including private ownership, entrepreneurship, competition, disjoint knowledge, and government failure—do not lend themselves to a doctrinal presumption against the welfare state, the best way to pay the government’s bills, and toppling terrible governments abroad—or, at least not to an extent like in other matters.
<table>
<thead>
<tr>
<th>#</th>
<th>Repeal or Liberalize or Phase Out or Eliminate</th>
</tr>
</thead>
<tbody>
<tr>
<td>8th</td>
<td>Repeal the expansion of money laundering laws held within the Patriot Act and the Bank Secrecy act of 1970 that require some businesses to keep tabs on customers and report activity to the federal government.</td>
</tr>
<tr>
<td>9th</td>
<td>Revisit Sarbanes-Oxley regulations.</td>
</tr>
<tr>
<td>10th</td>
<td>Liberalize the control or enforcement of equal opportunity/anti-discrimination in employment.</td>
</tr>
<tr>
<td>11th</td>
<td>Liberalize the SEC control of financial markets.</td>
</tr>
<tr>
<td>12th</td>
<td>Diminish union privileges.</td>
</tr>
<tr>
<td>13th</td>
<td>Reduce or abolish the minimum wage.</td>
</tr>
<tr>
<td>14th</td>
<td>Reduce FCC telecommunications regulations.</td>
</tr>
<tr>
<td>15th</td>
<td>Reduce federal aid to institutions of higher education and research programs.</td>
</tr>
<tr>
<td>16th</td>
<td>Liberalize Department of Homeland Security regulations on air-travel security.</td>
</tr>
<tr>
<td>17th</td>
<td>Decrease the Department of Energy’s subsidies for research and development in gas and oil industries.</td>
</tr>
<tr>
<td>18th</td>
<td>Reduce federal involvement of urban transit (subsidies and restrictions) in cities across the nation.</td>
</tr>
<tr>
<td>19th</td>
<td>Reduce the energy efficiency requirements and regulations put in place by the Department of Energy.</td>
</tr>
<tr>
<td>20th</td>
<td>Decrease the Federal Energy Regulatory Commission’s regulations on hydro-electric dams and interstate transmission of electricity and gas.</td>
</tr>
<tr>
<td>21st</td>
<td>Decrease the extent of health and safety regulations put in place by OSHA.</td>
</tr>
<tr>
<td>22nd</td>
<td>Liberalize banking regulations and therefore reduce the responsibilities of the Office of the Comptroller of Currency.</td>
</tr>
<tr>
<td>23rd</td>
<td>Liberalize immigration controls (let more people into the country).</td>
</tr>
<tr>
<td>24th</td>
<td>Liberalize the Consumer Product Safety Commission.</td>
</tr>
<tr>
<td>25th</td>
<td>Phase out regulations that force vehicle manufacturers to comply with fuel economy standards set by the Department of Transportation.</td>
</tr>
<tr>
<td>26th</td>
<td>Decrease federal involvement and regulation in the maritime shipping industry.</td>
</tr>
<tr>
<td>27th</td>
<td>Eliminate regulations pertaining to thrift institutions, which include savings banks and savings and loan associations.</td>
</tr>
<tr>
<td>28th</td>
<td>Liberalize FAA regulations on airline safety.</td>
</tr>
</tbody>
</table>
29th | Reduce regulation of marine fisheries by shifting to a system of individual fishing quotas (IFQs).
---|---
30th | Liberalize pollution and environmental controls/enforcement by the EPA.
31st | Liberalize food and safety controls by the Department of Agriculture.
32nd | Lighten the Federal Acquisition Regulations. These regulations deal with the way the government obtains goods and services as well as requirements on contractors for selling to the government.
33rd | Phase down the federal forest and rangeland research that deals with fisheries, endangered resources, invasive species, and other biological resources.
34th | Liberalize gun control policies.
35th | Decrease the Department of Transportation’s safety requirements for motor vehicles, along with a reduction in the Department’s regulation of anti-theft devices, promotion of safety devices and safety programs.

Of the 22 privatizations, the ranking emerging from the responses is shown in Table 2.

**Table 2: The Federal Privatizations Most Worth Discussing in the ERP, According to 11 GMU Economists**

<table>
<thead>
<tr>
<th>Rank</th>
<th>Privatization of Federal Government Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>Disaster insurance (make private and eliminate any subsidy)</td>
</tr>
<tr>
<td>2nd</td>
<td>U.S. Postal Service (along with liberalization of entry)</td>
</tr>
<tr>
<td>3rd</td>
<td>Amtrak</td>
</tr>
<tr>
<td>4th</td>
<td>Social Security accounts (even if only a small percentage of pay-ins)</td>
</tr>
<tr>
<td>5th</td>
<td>Space exploration (NASA)</td>
</tr>
<tr>
<td>6th</td>
<td>Power and electric infrastructure</td>
</tr>
<tr>
<td>7th</td>
<td>All job training and workforce assistance programs</td>
</tr>
<tr>
<td>8th</td>
<td>Disaster response agencies</td>
</tr>
<tr>
<td>9th</td>
<td>Water infrastructure (excluding natural channels and rivers)</td>
</tr>
<tr>
<td>10th</td>
<td>Federal loan programs</td>
</tr>
</tbody>
</table>
Management of the Bureau of Land Management’s public land (262 million acres of land and 300 million acres of subsurface mineral resources)

USDA Forest Service land (155 national forests containing 190 million acres)

Correction and prison systems

Federal Deposit Insurance Corporation

Air traffic control

Payroll processing

National Park Service land (79.3 million publicly owned acres which include national monuments, historic trails, lakeshores, recreation areas, and national parks)

Recycling centers

The provision of immunization services

Highways, parking, and bridges

Patent and Trademark Office

Federal libraries

APPENDIX 2:

APPENDIX 3:
Detailed breakdown of the JEL criticisms of the ERP. (Excel) Link.

REFERENCES


Klein, Daniel B. and Alexander Tabarrok. 2006. Quotations: Economists’ Judgments About the FDA. Is the FDA Safe and Effective?, an extensive website at FDAREview.org, hosted by the Independent Institute. Link


### ABOUT THE AUTHORS

**Daniel Klein** is professor of economics at George Mason University. He is coauthor with Alexander Tabarrok of the extensive website on the FDA ([FDAReview.org](http://FDAReview.org)) and with Adrian Moore and Binyam Reja of *Curb Rights: A Foundation for Free Enterprise in Urban Transit* (Brookings Institution), and co-editor with Fred Foldvary of *The Half-Life of Policy Rationales: How New Technology Affects Old Policy Issues* (New York University Press). He is the chief editor of *Econ Journal Watch*. His email address is dklein@gmu.edu.

**Michael Clark** is a Ph.D. student at George Mason University where he earned an MA in economics in 2006. His email address is mclark4@gmu.edu.
Benjamin Franklin and Colonial Money: A Reply to Michener and Wright—Yet Again

FARLEY GRUBB*

CONTINUATION OF THE EXCHANGE BETWEEN RONALD W. MICHENER AND ROBERT E. WRIGHT AND FARLEY GRUBB FROM THE MAY 2006 ISSUE OF EJW.

Michener and Wright Comment on Grubb (January 2006)
Grubb Reply (January 2006)
Michener and Wright Rejoinder (May 2006)

Leonato: Neighbours, you are tedious.
Dogberry: It pleases your worship to say so, …; but truly, for mine own part, if I were as tedious as a king, I could find it in my heart to bestow it all on your worship.
Leonato: All thy tediousness on me, ah?
—Shakespeare, Much Ado About Nothing, Act III, Scene V

MICHENER AND WRIGHT (2006C) REPEAT THEIR AD HOMINEM attack on me—repeat in that they offer little that is new or original. They simply rearrange and re-cloth their prior material and present it again.¹ Much of their rearrangement consists of rhetorical games—clever attempts

¹ Economics Department, University of Delaware, Newark, DE 19716.
The author thanks Catalin Cristian Popescu for research assistance and Tracy McQueen Cass for editorial assistance. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

¹ They not only repeat their material from Michener and Wright (2006a), but this repetition continues back through Michener and Wright (2006b, 2005) and Michener (1988, 1987). To be fair, they do occasionally present a new anecdotal quote or engage in some ersatz rhetorical jabs with my replies, but this is hardly enough to justify a new article.
to mislead the reader without being *de jure* dishonest. As in my past replies (Grubb 2005, 2006a), I will present new research—crafted here to reveal Michener and Wright’s attacks yet again for what they really are.

I will begin by deconstructing Michener and Wright’s (2006c, 251-253) opening rhetorical game, which will also serve to segue into the main body of the reply. There I will use the writings of Benjamin Franklin to evaluate Michener and Wright’s position on colonial money. With one possible exception, Franklin refutes Michener and Wright’s claims and supports my position on colonial money. As such, Michener and Wright’s endless attacks on me are really just endless attacks on Franklin. After that I will present new evidence and argumentation that decisively refutes Michener and Wright’s evidence that gold dominated the specie money supply in the middle colonies in the decade prior to the Revolution. I will close by briefly commenting on the appendix to their rejoinder.

**RHETORICAL GAMES (I)—THE OPENING ACT**

Michener and Wright (2006c, 251-253) open with two rhetorical games, the first to establish their motives and the second to establish the key point of disagreement between us—both cleverly designed to mislead the reader. They employ a cornucopia of rhetorical tricks including the bait and switch, the faux humility, the repetition ad nauseam, the self-evaluation congratulation, the outlier-as-average, and so on. But their preeminent trick is preemptively to accuse their opponent falsely of what they themselves are actually doing in order to insulate themselves from criticism for doing said behavior. The deconstruction of their two opening rhetorical games, besides providing illumination and some amusement, will set up the substantive portion of the reply.

**Game One: Michener and Wright’s First Paragraph**

Michener and Wright’s first paragraph (2006c, 251) opens by asserting that Grubb is engaged in a radical rewriting of history. Next they express humility in the face of the complexity of the colonial economy so the reader will see them as not being ideologues or curmudgeons. They follow this by claiming that they have no codified position on colonial
money and so have nothing at stake (no ulterior motives) other than to alert scholars to Grubb’s questionable research. It is a wonderfully artful paragraph.

Now let’s deconstruct it, peeling it like an onion backwards from the outside in. Their last sentence asserts that their only motive is to alert scholars to Grubb’s questionable research. But wait a minute. All the assumptions, caveats, references to alternative possibilities, etc., that are a part of applied quantitative research are already well laid out for the reader to assess in Grubb (2004, 2006b). Nothing is hidden there. So that cannot be Michener and Wright’s true motive. So what are they really up to?

Let’s leap to the first sentence of the paragraph and deconstruct it—which will help answer the question just asked. Michener and Wright assert that Grubb is radically rewriting history. This is an example of the rhetorical trick—repetition ad nauseam. It is a claim they repeat every time they write about Grubb (Michener and Wright 2005, 2006a). They present no research to directly evaluate or support this claim, nor do they address the evidence to the contrary (e.g. in Grubb 2005, 2006a). Instead, they hope that the mere force of repetition will plant this idea in the mind of the reader. The impression they hope to leave is that Grubb’s rewriting of history is so outlandish and so far outside the mainstream that it must be highly suspect and so should be discarded or not even read.

Their assertion here is also an illustration of their most frequently used rhetorical trick—preemptively accusing one’s opponent falsely of what oneself is doing to try to insulate oneself from criticism of doing exactly that. That Grubb is rewriting early American history in some outlandishly radical way is true only in the minds of Michener and Wright and true only in comparison to Michener and Wright’s position on colonial money. As will be illustrated below, Grubb’s position on colonial money is squarely within that taken by Benjamin Franklin—about as mainstream as one can get.

In truth, it is Michener and Wright’s position on colonial money that is outside the scholarly mainstream. No one working within the field of colonial money and colonial history seriously believes their position. For example, Bruce Smith (2000), an eminent macro-economist who published extensively on colonial money, in an e-mail sent to me shortly before he died of brain cancer summed up Michener’s position on colonial money as, “His interpretation of events is accepted by him [Michener] and Ben

---

2 If Michener and Wright made their solicited and unsolicited correspondence to journal editors, grant agencies, and authors on these matters available, the reader would blanch. A different motive would appear.
McCallum. To my knowledge, no one else. I think that virtually nothing in the Michener [(1987)] Carnegie-Rochester paper is correct. . . . John McCusker supported me on this.”

Now let’s move into the heart of the onion and deconstruct what is left of Michener and Wright’s first paragraph. They say that they “are humbled by the complexity of the early American monetary system and the meagerness of the available evidence.” This [faux] humility makes them sound like tempered and reasonable scholars. But wait a minute. If the colonial economy is really that complex, too complex for them to comprehend, explain, or model for us—as they frequently say when cornered, and if the evidence is that meager, then how can they engage in such strident attacks with such certitude against Grubb, as well as against other scholars working on colonial money? They cannot have it both ways—being humbly ignorant as well as all-knowing arbitrators of truth in the field.

The real point of their “complexity” and “meagerness” of evidence sentence, apart from the faux humility rhetoric, is to insulate their model from any meaningful test and so avoid having it falsified. Their position is so complex, so unexplainable, and so unfathomable that it is irrefutable. No evidence or arguments to the contrary can touch them. This leads to the last part of their opening paragraph.

In the middle of their first paragraph Michener and Wright (2006c, 251) appear to claim that they no longer hold that specie was generally plentiful, that exchange rates were meaningfully fixed, and that bills of credit often circulated as a media of exchange among the populace of neighboring colonies. Seemingly any outcome here is possible. They say this so the reader will see them as not being dogmatic ideologues and so see them as disinterested evaluators of Grubb’s research. But wait a minute—what they just said sounds a lot like Grubb’s position. Why then do they make such a fuss? While seemingly any monetary outcome is possible, they still claim to know one thing with certainty and that is whatever Grubb says is absolutely wrong and should be dismissed, not used, and not even read. Now there is an ad hominem position if ever there was one!

Have they really had a change of heart on the issues of specie scarcity, fixed exchange rates, and currency substitution, or are they playing a rhetorical game on the reader? That they really have abandoned their prior position (Michener 1987, 1988; Michener and Wright 2005, 2006a, 2006b) is inconsistent with the rest of their rejoinder and its unwavering attack on Grubb and certitude over his errors. The syntax in their rejoinder is most consistent with their either hoping that the reader will not refer back to
their prior works when reading their rejoinder to note the contradictions or with them not fully comprehending their own model.

They phrase their faux retraction in terms of not having made “blanket” claims, namely absolute, always, and forever statements. But of course no one accuses them of such. The debate is not over absolutes but over what is dominant, average, typical, most likely. In other words, they use the existence of atypical exceptions to make it appear as if they do not hold the radical positions that they do. For example, they, along with most scholars, accept that specie was scarce in Massachusetts in the 1740s, New York bills of credit did not circulate in Georgia, and South Carolina exchange rates rose dramatically in the 1710s and 1720s. But none of these exceptions are for the time and place of our debate, namely for the middle colonies from New York to Maryland during the period of the 1720s to the Revolution. For that time and place, Michener and Wright still believe as a general rule that the money supply was dominated by specie, exchange rates were fixed enough to foster a high degree of currency substitution, and so bills of credits circulated as a media of exchange among the general populace of neighboring colonies. In other words, they really do have a lot at stake because Grubb’s findings (2004) are inconsistent with their claims. Their motives are not as pure as they wish the reader to think.

The claim in their first paragraph (Michener and Wright 2006c, 251) that exchange rates oscillated within broad specie points has a slightly different purpose. Such a claim is still a claim that exchange rates were fixed, even though no colonial government ever actively intervened in the foreign exchange market to manage their exchange rate in an on-going manner. They also appeal now to legal tender laws as the primary mechanism that fixed exchange rate rather than to some mysterious conspiracy of merchants. As such, they have not set aside their assumption of fixed exchange rates for Grubb’s position that rates were more floating than fixed.

Their claim that “exchange rates oscillated within broad specie points” is merely an effort to escape a falsifiable position, i.e. to insulate their model from meaningful testing. In other words, there is no movement in exchange rates that can refute their assertion that exchange rates were fixed. Exchange rates can fluctuate widely and wildly and still Michener and Wright’s model of fixed exchange rate regimes in colonial America is upheld (Michener 1987). In other words, their model is irrefutable.\(^3\)

---

\(^3\) On the next page they say that “‘Dollar’ and 7 shillings 6 pence were synonymous in late colonial Pennsylvania…” (Michener and Wright 2006c, 252) Apparently, the oscillation of
However, if rates can fluctuate that much then over the relevant range you have a de facto floating exchange rate regime. And the whole point of Michener’s (1987, 1988) assertion (carried through Michener and Wright) of fixed exchange rates is to establish tight enough (costless enough) currency substitution that colonies’ money supplies are determined principally by their money demands. In other words, specie and the bills of credit of the various colonies all flowed more-or-less costlessly across colonial borders to accommodate given money demands. Thus, prices in a given colony will not correlate with any one component of its money stock, which is the key payoff point of their modeling exercise (Grubb 2006a, Michener 1987). Now, if exchange rates oscillate within a “broad” range, Michener and Wright’s low-cost currency-substitution model is destroyed and their entire edifice of how the colonial monetary system worked falls apart. Either they do not understand their own model of colonial money or they hope the reader will not refer back to their prior work when reading their attack on Grubb here.

Having been deconstructed, Michener and Wright’s first paragraph (2006c, 251) can now be re-imagined as saying: *We have a position on colonial money that no one seriously believes in and it pisses us off. How dare anyone disagree with us? Such people are wrong-headed and misguided. After all, our position is irrefutable. Our model of colonial money is so complex that it cannot be explained by us nor can it be comprehended and tested by anyone else. There is no way our model can be falsified, especially given our ability to dispute any evidence. Our motive here is to hunt down unbelievers like Grubb and drive them from our temple of colonial money.*

**Game Two: Their Second and Third Paragraphs**

Michener and Wright in their next two paragraphs (2006c, 252) seek to establish the key point of disagreement between themselves and Grubb. That being, according to them, Grubb confuses units of account with media of exchange while they make no such confusion. This makes all of Grubb’s research wrong and so all of it should be discarded and not even read. This is the key point that they repeat ad nauseam in every paper they have written attacking Grubb, and even in papers not directly dealing with exchange rates “within broad specie points” only lasts one page for them. There claim here is also an example where they confuse the customary exchange rate used when referring to units of account with the actual exchange rate used when trading in different physical mediums of exchange.
Grubb (Michener and Wright 2005; 2006a; 2006b, 24). They need to establish this point because without it they have no paper, i.e. they have no meaningful position with which to attack Grubb and others working on colonial money who found results that conflict with their position. Yet, their presentation here is a rhetorical game designed to cleverly mislead the reader without being de jure dishonest.

Let’s read the language in their second paragraph carefully. In reference to evidence enumerated in Pennsylvania pounds, Michener and Wright say “We believe that…[it] usually meant…” units of account. They say “usually meant” which means not always meant. So Michener and Wright are saying that evidence enumerated in Pennsylvania pounds sometimes meant media of exchange (Pennsylvania paper bills of credit) and sometimes meant just a unit of account. Now let’s see how Michener and Wright characterize Grubb’s position. They say that “Grubb believes that in certain transactions…” evidence enumerated in Pennsylvania pounds “invariably meant...” media of exchange. They say “certain transactions” not all transactions. So Grubb must believe that evidence enumerated in Pennsylvania pounds sometimes meant media of exchange and sometimes meant just a unit of account. [This interpretation of Grubb also accords with, and is well stated in, Grubb (2004, 331-332).]

Given this closer look at their characterizations of their and Grubb’s position on this issue, darned if I can see any difference. Both Michener and Wright and Grubb hold that evidence enumerated in Pennsylvania pounds sometimes meant units of account and sometimes meant media of exchange. There is no conflation of units of account with media of exchange by Grubb or by Michener and Wright as a general principal. They both believe the same thing. As such, Michener and Wright have continuously mischaracterized their key point of disagreement with Grubb, and they have purposely done so because revealing the true point of disagreement would kill their attack.

Since both Grubb and Michener and Wright believe the same thing, namely that evidence enumerated in Pennsylvania pounds can mean either media of exchange or units of account, the key difference between them becomes what criteria they use to decide when evidence is most likely reflecting units of account as opposed to most likely reflecting media of exchange. Michener and Wright never formally state their criteria. In fact they say it is just too complex to comprehend and explain to the reader. But they obviously employ some criteria, and if one works back through their publications one consistent criterion comes through, namely that what ever is inconsistent with their model of colonial money is branded as unit-of-
account evidence and dismissed. Evidence which supports their model is more likely to receive their approval as media of exchange. This is a purely self-serving selection criterion, serving to make their model irrefutable. No evidence can ever be uncovered that will refute their model because such evidence by their criterion can be stamped as just unit-of-account evidence and so dismissed.

By contrast, Grubb (2004, 2006b) seeks to establish some independent criteria, using objective standards, for assessing when quantifiable evidence would most likely reflect media of exchange rather than just units of account. The outcome of this search led to evidence which contradicts many aspects of Michener and Wright's model of colonial money. This in turn is what has put them into such an apoplexy.

The real point of disagreement between Grubb and Michener and Wright is not that Grubb is somehow naively and hopelessly confused over media of exchange versus units of account, but is over the models of colonial money each bring to the table. So let’s remind ourselves what those differences are, and let’s start with Grubb’s model. For the middle colonies between the 1720s and Revolution, Grubb (2004) finds that as a general rule the paper money of a given colony (when it issued paper money) dominated monetary transactions within that colony among its general populace. Bills of credit of neighboring colonies and specie were used more rarely. Specie became less scarce among these transactions in the decade before the Revolution as relatively more Spanish silver dollars flowed into the colonies through trade surpluses and lingered longer before flowing back out to cover trade deficits. He finds, as a general rule, little evidence of large scale or tight currency substitution between specie and paper money and between the paper monies of neighboring colonies with regard to usage among the general populace. As such, his evidence is consistent with the institutional evidence indicating that the colonies operated under more of a floating than a fixed exchange rate regime.5

---

4 He also assumes that when a unit of account has no physical representation, people behave in a certain way and mean a certain thing when referring to it. But when a physical representation is created for that unit of account—a physical bill of credit is put into circulation—people will behave differently and mean something different when referring to its value than when there was no physical representation of that unit of account. A change in constraints elicits a change in behavior.

5 Among merchants engaged in foreign, overseas, and cross-colony trade specie and the paper money of other colonies may well have been important media of exchange within their monetary transactions. Certainly, a lot of specie passed through the colonies, imports then exports of specie, via these merchants dealing in the extensive multi-lateral foreign trade of the day. And certainly merchants dealing in cross-colony trade engaged in the
There was a strong “inside” versus “outside” dichotomy in colonial monetary systems.

By contrast, Michener (1987, 1988) and Michener and Wright (2005, 2006a, 2006b, 2006c) postulate that these colonies operated under a fixed exchange rate regime. Rates were held fixed by legal tender laws and/or a conspiracy of colluding merchants, so proven by the existence of several coin rating tables. They further postulate that these fixed exchange rates produced tight enough (costless enough) currency substitution parameters that money supplies were indeterminate. As such, specie and the bills of credit of neighboring colonies more-or-less flowed costlessly across colonial borders to accommodate money demand. There was no “inside” versus “outside” dichotomy to colonial monetary systems. While cross-colony circulation of paper money was frequent and substantial, they hold that specie still dominated the media of exchange among monetized transactions throughout—else currency substitution of paper for specie money would have caused substantial inflation which was not observed (see also the discussion in Grubb (2004, 2006a)).

Finally, they claim that gold, not silver, dominated the specie money supply in the middle colonies in the decade prior to the Revolution.

6 They also argue that when a unit of account has no physical representation, people behave in the exact same way when referring to it as when a physical representation is created for that unit of account, i.e. when a physical bill of credit is put into circulation—as such for the Pennsylvania pound before versus after 1723 (Michener and Wright 2006a, 6-8; 2006b, 25; 2006c, 252-253). As such, their position is that a change in constraints elicits no change in behavior. This is a very odd position for a Chicago economist to take. This is also part of their bait and switch rhetorical tactic, i.e. confusing evidence from before bills of credit were issued with behaviors after bills of credit were issued.

7 Their model and arguments also make no distinction between what circulates as a media of exchange among the general populace and what is used by merchants engaged in foreign and cross-colony trade.
BENJAMIN FRANKLIN: WHAT WOULD HE THINK ABOUT THIS?

No other American was involved over as long a period of time with so many different facets of colonial paper money as was Benjamin Franklin—certainly no other American with such a preeminent stature in science, statesmanship, and letters. Franklin arrived in Philadelphia the year paper money was first issued by Pennsylvania (1723). And he was the elder statesman at the U.S. Constitutional Convention in 1787 that brought this colonial monetary system to an end (Grubb 2006c). In between Franklin was a keen observer of, and commented often on, the colonial monetary economy. He wrote pamphlets, treatises, and commentaries on paper money. He designed and printed paper money for various colonies. He entertained ideas about and proposed alternative monetary systems. As an assemblyman for the colony of Pennsylvania he was involved in the debates over and management of that colony’s paper money. And as a lobbyist for various colonies to the British court, he dealt with conflicts over colonial paper money that arose between Britain and her colonies. As such, Franklin’s views should carry considerable weight in determining what the colonial monetary system was like.

On virtually every aspect of colonial money, Michener and Wright’s model is contradicted by, while Grubb’s model is consistent with, Franklin’s views. Franklin sees specie as being relatively scarce as a media of exchange among the general masses in the colonies. Even within the colonial merchant community involved in overseas trade, Franklin sees specie availability as subject to large swings due to ubiquitous foreign trade shocks that frequently stripped a colony of its specie. Fluctuations in the amount of specie and its relative scarcity in the colonies were due to foreign trade shocks, chronic balance of trade deficits, and British-imposed absences of trade and capital controls in the colonies. It was not due to currency substitution between issuances of paper money and specie. Franklin sees colonial paper money as predominantly an inside money, seldom being exported or circulating as a media of exchange among the general populace of neighboring colonies and states. Its value is governed by the quantity of

---

it issued relative to the demands of internal commerce and not by the operation of some fixed exchange rate regime. Finally, Franklin explains that Pennsylvania’s legal tender laws did not create a fixed exchange rate regime. Legal tender laws served a different purpose, not unlike that served in today’s economy.

In 1723 at age 17 Franklin travels from Boston to Philadelphia. He has specie in his pocket and no paper money, even though he says that the primary money used in Boston at this time was paper money. He returns to Boston in 1724 again with only specie in his pocket and no paper money even though he says that prior to Pennsylvania issuing paper money in 1723 the colony had been stripped of its gold and silver. On his return to Philadelphia, Franklin is instructed to collect a debt of 35 pounds currency due to be paid in Pennsylvania to a friend of Franklin’s brother who lives in Rhode Island. He is to collect the money in Pennsylvania and keep it until he has “directions what to remit it in” to Rhode Island. Franklin travels to London from Philadelphia late in 1724 and again has only specie in his pocket and no paper money. (Labaree, Willcox, and Oberg 1959-1998, 1:147, 152, 14:78; Smyth 1907, 1:254, 260-262, 276)

Franklin’s behavior in the 1720s is illustrative of the fact that a colony’s paper money served predominantly as an inside money and did not readily cross colonial borders as a media of exchange among the general populace. Specie, though relatively scarce, was the principal cross-colony and cross-oceanic media of exchange. Franklin’s writings throughout his life are consistent with his 1720s behavior. In 1729 he says, “…we [Pennsylvanians] have already parted with our silver and gold…” in trade with England, and that the difference between the value of paper and that of silver is due to “the scarcity of the latter.” In 1752 as a Pennsylvania Assemblyman he reports that paper money furnished “…the country with a medium of trade, and of a kind that could not, to any purpose, be exported…” And that without paper money “…the province will then be left without any currency, except that precarious one of silver, which cannot be depended on, being continually wanted to ship home, as returns, to pay for the manufactures of Great Britain.” Similarly, in 1757, he reports that the specie inflows during the late war were “…since drawn out of the province for payment of the army…and for payment of our debts to merchants in England; so that a piece of gold is now rarely received in payment.” (Labaree, Willcox, and Oberg 1959-1998, 1:147, 152; 4:345, 348; 7:123)

These observations continue after Franklin goes to London as a lobbyist for the colonies. In 1676 he writes that Britain could retain specie
as a media of exchange because Britain has capital and trade controls. But the colonies are not allowed such controls and so “…they cannot keep gold and silver among them sufficient for the purposes of their internal commerce…”, and that “Paper bills…have therefore in the colonies long been substituted for real money.” Franklin also argues that the claim so often made that it is the very issuance of paper money itself that causes specie to be carried out of the province and so is the primary cause of specie scarcity (the currency-substitution claim) is “…a mere speculative opinion, not founded on fact in any of the colonies. The truth is, that the balance of their trade with Britain being generally against them, the Gold and Silver is drawn out to pay that balance; and then the necessity of some medium of trade has induced the making of paper money, which could not be carried away….The balance of trade carry’d out the gold and silver as fast as it was brought in…” Franklin also notes that specie has a “universal estimation” but “…that very universal estimation is an inconvenience which paper money is free from, since it [specie] tends to deprive a country of even the quantity of currency that should be retain’d as a necessary instrument of its internal commerce; and obliges it to be continually on its guard in making and executing at a great expense the laws that are to prevent the trade which exports it [specie].” Finally, even into the early 1780s, Franklin continues to explain that, “It has been long & often observed, that when the current money of a country is augmented beyond the occasion for money, as a medium of commerce, its value as money diminishes…if some means are not found to take off the surplus quantity. Silver may be carried out of the country that produces it, into other countries,… Paper money not being easily receiv’d out of the country that makes it, if the quantity becomes excessive, the depreciation is quicker & greater.”

That Franklin saw a colony’s paper money as predominantly an inside money is also revealed in the nature of his grand scheme for supplying the colonies with a general paper currency. In 1765 Franklin proposed that the British establish a North-American-wide land bank (supervised in part by the Bank of England) where paper money would be issued upon the security of mortgaged land in any colony. This paper money could then be

---

9 In 1760 the British Board of Trade reached the same conclusion. In reference to Pennsylvania they said, “…that the paper currency…has been issued in this colony, and in the other provinces of North America,…to serve as a medium of circulation within the province; the balance of trade being so much against them, that gold or silver is very difficult to be procured.” [italics added] (Labaree, Willcox, and Oberg 1959-1998, 9: 147)
made a legal tender across all the colonies (this may be the first proposal ever made for what amounts to a “national” paper currency for North America). Franklin’s proposal presupposes that the current system of separate colony-specific issuances of paper money functioned predominantly as inside money, not generally flowing as a media of exchange across colonial borders among the general populace. His proposal, among other things, would overcome this “localism” of paper money and yield a more “national” usage. Alas, the uproar over the Stamp Act led Franklin to bury, hide, and disavow his proposal. It was forgotten. (Labaree, Willcox, and Oberg 1959-1998, 12: 51-61)

The one exception Franklin relates to his general view that a given colony’s paper money was predominantly an inside money is for Maryland in the early to mid-1760s. This case is also used by Michener and Wright (2006a, 24-27) as evidence for their claim that cross-colony circulation of paper money was commonplace. Early in 1767 while in London Franklin wrote up some “Hints of Arguments” regarding legal tender laws to be used in his effort to get the British to repeal their last currency-act legislation. There he says that as Maryland moved towards retiring all its outstanding paper money in 1764, that money ceased to serve as a medium of exchange. As a consequence, having no paper money of their own and insufficient specie to transact commerce, Marylanders acquired and used Pennsylvania paper money as a medium of exchange. [New Maryland paper money to replace that retired in 1764 would not be issued until 1767.] (Labaree, Willcox, and Oberg 1959-1998, 14: 37-38; Newman 1997, 164-167)

This one exception hardly invalidates Franklin’s general rule that a colony’s paper money functioned predominantly as an inside money and seldom circulated as a media of exchange among the general populace of other colonies (especially in other colonies that had their own paper money). In addition, the extent of this exception remains questionable. Was it just some minor increase in the usage of Pennsylvania pounds for a few years by merchants dealing in cross-colony trade or did Pennsylvania pounds as a media of exchange penetrate the general populace of Maryland over some extended period for use in their internal trade? The little evidence out there speaks to the former and not the latter (see Grubb 2006a, 63-64; Michener and Wright 2006a, 24-27). Finally, doubts about the extent of this exception can also be taken from the fact that Franklin publishes all the arguments out of his “Hints of Arguments” except the story of Marylanders using Pennsylvania pounds. Why was this one argument not published? Was it just an oversight? Or was it because this story was not generally true, or ceased to be true after a year or two, or was
exaggerated and people in America would know it? (Labaree, Willcox, and Oberg 1959-1998, 14: 32-39, 77-87)

Finally, Franklin writes a lot about the role and function of legal tender laws—a key point of difference between Grubb's model and Michener and Wright's model of colonial money. For Pennsylvania, Franklin contradicts Michener and Wright's assertion that legal tender laws effectively fixed exchange rates and thus the value of paper money. Franklin argues that it is the quantity of paper money relative to the volume of internal trade within a colony that governs its value, and not legal tender laws. Exchange rates between paper money and specie are determined by their relative supplies or scarcities compared with the relative demands for inside versus outside trade. For example, Franklin notes that, “…too great [a] quantity [of paper money] has, in some, colonies, occasioned a real depreciation of these bills, tho’ made a legal tender…”

For legal tender laws to effectively enforce a fixed exchange rate, the courts must impose that fixed exchange rate when settling cases when the legal tender is substituted for some other contracted payment. Just like for our modern economy, Franklin notes that this did not happen. Courts used the expected market-determined exchange rate and not some fixed rate to equate values between media of exchange when settling cases. In 1767 Franklin observed, “…it having ever been a constant rule there [in the middle colonies] to consider British debts as payable in Britain, and not to be discharged but by as much paper [money] (whatever might be the rate of exchange) as would purchase a bill for the full sterling sum.” Franklin goes on to write in a draft petition, “…in the Courts of Justice [in Pennsylvania], full satisfaction has ever been given in discharge of debts due to the British merchant…”

In 1760 the British Board of Trade made the same observation. They concluded that Pennsylvania had been exempted from the Currency Act of 1751 that forbid making paper money a legal tender because “…the province had, without a Law, come of itself very near the regulation which the Law would have prescribed.” (Labaree, Willcox, and Oberg 1959-1998, 9:149; 14: 34-36, 80, 185)

Rather than fixing exchange rates, Franklin explains that legal tender laws had a different function. Because colonial paper money was backed by a government pledge of redemption at some future date, there could be a
time discount built into its current value. Just as a commercial bill of exchange trades at a discount off its face value reflecting, in part, the time to its redemption, colonial paper money might suffer some discount off its face value reflecting the time to its redemption. Legal tender laws simply removed this time discount dimension from paper money and so caused it to trade at its nominal face value, other things equal. In other words, unlike with comparable commercial bills of exchange, one Pennsylvania pound redeemable in 2 years and one Pennsylvania pound redeemable in 5 years would both trade at their face value to each other concurrently under a legal tender law. As such, legal tender laws served merely to remove confusion over the internal nominal valuation of a colony’s various paper money issues. (Labaree, Willcox, and Oberg 1959-1998, 14: 35-39, 86)

The other purpose of legal tender laws was “…the convenience to the possessor where every one is oblig’d to take them…” (Labaree, Willcox, and Oberg 1959-1998, 14: 34) While not fully articulated by Franklin, from the totality of his writings on the subject, this appears to mean solving the short-run or temporary hold-up problem in trade. Franklin sees outside monies (specie) as being prone to substantial and unexpected short-run fluctuations in availability, say due to unexpected foreign trade shocks. Inside monies (the paper monies of the colonies) were not prone to these effects. If a creditor who was owed a payment in outside money could refuse being paid in an equivalent amount of inside money, say at its expected market exchange rate of outside to inside money, then the creditor could exert undue leverage over the debtor when outside money was unexpectedly and temporarily scarce. Under such a circumstance, the creditor could extract more payment than the expected market equivalent of the outside money owed by threatening the debtor with debtors prison for non-payment of the specific outside money. Legal tender laws removed this creditor-extortion possibility—Shylock’s demand for a literal pound of flesh and no equivalent (Shakespeare, Merchant of Venice).

In conclusion, on every significant point of difference regarding colonial money, Franklin supports Grubb and refutes Michener and Wright. As such, Michener and Wright’s endless attacks on Grubb are really just endless attacks on Franklin. Now all this evidence is anecdotal, and I am the first to put little definitive weight on anecdotal evidence, though there is a fair amount here and Franklin is remembered as ‘the wisest American’ of his generation (Spiegel 1987, 415). Michener and Wright could always brand Franklin as hopelessly partisan on monetary issues and so dismiss him out of hand. But Michener and Wright live and die by unparsed anecdotal evidence. So by their own methodological standards the evidence here
utterly demolishes their model of colonial money and strongly supports that used by Grubb (2004). As such, Grubb is well within the mainstream and Michener and Wright are far outside the mainstream on the issue of how the colonial monetary system worked.

“HORTENSIO PEACE! THOU KNOW’ST NOT GOLD’S EFFECT”
—Shakespeare, *The Taming of the Shrew*, Act I, Scene II

“[SILVER] DOLLARS ARE HERE NOW AS PLENTY AS DIRT”
—Johnson (1957, 12: 745)

Michener (1987, 1988) and Michener and Wright (2005, 2006a, 2006b, 2006c) claim that specie dominated the colonial money supply. In addition, they claim (2006a, 2006c) that in the decade prior to the Revolution this specie was primarily gold. This last claim will surprise scholars of colonial America who will likely find it preposterous.

Michener and Wright (2006a; 2006c, 12) support their claim by saying “Gresham’s Law” (bad gold coins drove out good silver coins) and by presenting one new anecdotal quote that appears to indicate that silver was scarce relative to gold. However, Michener and Wright’s use of Gresham’s Law is just theoretical bluster. On closer examination it is not applicable, i.e. they fail as economists here. And their new anecdotal quote is an evidential bluff. As with much of their anecdotal evidence, it is presented in a manner that misleads the reader, i.e. they fail as historians here.

The Theoretical Bluster: Gresham’s Law

Gresham’s Law—bad money drives out good money—when applied to gold versus silver coins, depends on someone fixing the exchange rate of gold to silver at a rate different than that determined in the marketplace by the forces of supply and demand. Typically this is done by a government through force of law at that government’s mint (Greenfield and Rockoff 1995, Harris 1987, Rolnick and Weber 1986). But in this period there are no mints in British colonial America and no British colonial governments are

---

11 In the colonies and for the periods that are the focus of their attacks on Grubb.
fixing and enforcing gold to silver exchange rates different from the rates in the open market. Therefore, Gresham’s Law is not applicable here. As such, Michener and Wright (2006a, 2006c) do not know what they are talking about.

Now, in an effort to salvage their Gresham’s Law approach, Michener and Wright could be interpreted as arguing that a private cartel of merchants in colonial America operated effectively in place of a government and its mint in fixing the exchange rate between gold and silver in America to be different than the rate determined in the open market. They certainly appear to argue that way. They constantly present coin-rating tables which they take as prima facie evidence of this, and then deduce from other evidence that this cartel must have been fixing this rate at some non-market determined level (Michener 1987, 1988; Michener and Wright 2005, 2006a, 2006b, 2006c).12

Michener and Wright’s interpretation, however, violates rational maximization. If some informal cartel of colonial merchants agreed to over-rate gold relative to silver, then other colonial merchants would stick these cartel merchants with gold instead of silver when trading with them. Gold would pile up among these cartel merchants and silver would come to dominate specie circulation in the rest of the colonial economy outside the cartel: the exact opposite result from what Michener and Wright claim would happen. Finally, the merchant cartel would lose money, being stuck with a lot of gold that was valued less in the international market than what they valued it at when trading for it.

Now Michener and Wright (2006a, 31-33) do estimate that the relative price of gold to silver in colonial America was higher than in England in the decade before the Revolution, i.e. \((Pg/Ps)a > (Pg/Ps)e\) \([P = \text{price; } g = \text{gold; } s = \text{silver; } a = \text{colonial Pennsylvania; } e = \text{England}]\). Thus, silver, relative to gold, should flow from colonial Pennsylvania to England in this period. But this isn’t Gresham’s Law. It is just market arbitrage, and is a necessary condition in the face of the relative rise in the amounts of silver compared with gold being produced in the Americas. Far from contradicting Grubb’s position on the rise in the use of silver relative to gold in Pennsylvania in this period, their evidence is actually consist with Grubb’s findings (2004, 340).

---

12 This coin-rating-table evidence is not new to Michener and Wright. They did not discover it. See Solomon (1976, 35, 39, 41). But neither is it interpreted by other scholars as Michener and Wright do.
Commodities are flowing all over the Atlantic economy as the forces of market arbitrage react to shifting relative price inequalities between locations. Relative prices in receiving locations must be higher, and remain higher, than in sending locations by at least the cost of transportation to sustain an on-going flow of commodities across the Atlantic. Specie is no exception. The New World produced the lion’s share of global specie and this specie had to be distributed through trade to the rest of the world. This was a long-run on-going flow event, not some fixed stock position in an instantly arbitrated static equilibrium as Michener and Wright (2006a, 31) claim.

Barrett (1990, 224-226, 238) finds that “…from 1493 to 1800, 85 percent of the world’s silver and over 70 percent of its gold came from the Americas. In each century both absolute production and the relative share of the Americas in the world increased…” The production of gold relative to silver, however, was highly concentrated. “Production of gold in Brazil [Portuguese gold] lasted only slightly more than half a century, from before 1700 to the 1750s…” As gold production plummeted after 1750 silver production continued to rise rapidly. Between 1726-1750 and 1751-1775 Barrett (1990, 242-243) finds that New World silver production increased 26 percent, and while silver exports to Europe increased they did not keep pace with this expansion. As such, the retention of silver in the New World (new production minus exports) increased by 53 percent. These observations accord well with what Grubb (2004, 340-341) finds for Pennsylvania.

Now in the decade before the American Revolution, this burgeoning production of silver relative to gold had to be moved from the New World to the Old. The British North American colonies were an important conduit in this movement—trading goods to Spanish America for silver and then exporting this silver to England for English goods. For this to happen, silver had to be under-valued relative to gold in the Americas compared with that in Europe. [Gold had to be over-valued relative to silver in the Americas compared with that in Europe.] In other words, \((P_g/P_s)a > (P_g/P_s)e\) which is exactly what Michener and Wright report. Michener and Wright have merely discovered the conditions necessary for global market arbitrage in an ongoing dynamic setting. For them to use this observation to imply that little silver was left in the middle colonies for use as a media of exchange relative to gold because it was all exported to Europe is to confuse a fixed static setting for an ongoing dynamic one.

---

13 See also Attman (1986, 26-29, 66, 70) and Fischer (1996, 128-129).
Not only is Michener and Wright’s observation of \((Pg/Ps)a > (Pg/Ps)e\) a sign or necessary condition of, and caused by, the relative excess of silver (or access to silver) in the middle colonies, but given the rising retention of silver in the Americas in this period cited above (Barrett 1990, 242-243), this relative gold-to-silver price inequality was not sufficient to carry all the increased silver production (relative to gold) to Europe. As such, Michener and Wright’s analysis is inconsistent with their conclusion and actually supports Grubb’s (2004) position.

The Evidential Bluff: Quoting from the Papers of Sir William Johnson

In support of their claim that gold, not silver, dominated the specie money supply in the middle colonies in the decade before the Revolution, Michener and Wright (2006c, 12) present one new anecdotal quote as proof of their position. They say,

We recently unearthed evidence that supports our view that Spanish dollars were relatively scarce in Pennsylvania during this period. In November 1768 Sir William Johnson enlisted James Tilghman of Philadelphia to gather some for him. Tilghman sought to accommodate him, but cautioned “its very probable we may be obliged to make up some deficiency with half Jo[hannes]s, as the dollars are extremely scarce” (Tilghman 1768). Tilghman’s comment is perfectly sensible under our interpretation, but problematic for Grubb, since Grubb (2004, 340, Fig. 1) indicates that dollars, instead of being “extremely scarce” in Pennsylvania in 1768, were more plentiful than they had been at any time in the previous four decades.

But as is often the case with the anecdotal evidence that Michener and Wright present, they have misled the reader and are completely wrong about what their historical source really says about the relative amount of silver to gold in this period. The Papers of Sir William Johnson (1921-1965), 14 well-indexed volumes, are in your nearby university library. They are not hidden or buried or in need of being “unearthed” as Michener and Wright claim to have done. And when the above quote presented by Michener and Wright (2006b, 12) is presented in just a slightly expanded version, a very

ECON JOURNAL WATCH 502
different impression emerges. Tilghman actually wrote (Johnson 1928, 4: 507-508),

We are collecting dollars with all possible dispatch in order to fulfill our Engagements on Account of the Indian purchase and I am in great hopes we shall nearly accomplish our Quantity in time however you may depend upon having the full Sum in time tho its very probable we maybe obliged to Make up some deficiency with half Jo$. as the dollars are extremely scarce.

Based on this fuller quotation, silver dollars appear to dominate the specie money supply with at best a few gold coins around—enough to “make up some deficiency” when pressed for time in the hasty fulfillment of a large order for silver dollars. In no way does this evidence mean what Michener and Wright try to make it mean (that as a general condition silver is very scarce and gold is very plentiful), nor is this evidence at odds with the evidence in Grubb (2004, 340).

In addition, when the complete works of Johnson (1921-1965) are consulted only a handful of references to gold Johannes can be found compared with an overabundance of references to silver dollars. And lest Michener and Wright try to play their game of asserting that any reference to “dollars” are just unit-of-account references and not references to actual coins in order to dismiss any contrary evidence to their position, let’s quote a few more passages. Abraham Mortier wrote Johnson in November of 1768 (Johnson 1928, 6: 458),

I wrote you in September last before I set out for Fort stanwix requesting you would send me a Parcel of Dollars for to pay the Indians with,... I had reason to Expect you would answer my Demand, but I never had even a Line from you on the Subject, which distressed me greatly and Obliged me to set every Engine to work on my own Credit to get between four and five thousand pounds in Dollars, nay I was obliged to Borrow 3000 Dollars from the mohawks money which Mr. Remsen then paid them for Kayadarosseras—

Allan Grant wrote Johnson in August of 1769 (Johnson 1957, 12: 745),
As to the report of the Traders here saying that you gave pewter Dollars to the Indians there has been no such thing said by any body here, the Traders know very well to the Contrary, they have all made fortunes here this Summer entirely owing to the Number of Dollars you gave the five Nations this Summer, they are much better off than the Commanding Officer this year. Some Cayugas asked the traders here if they were not pewter Dollars, I fancy owing entirely to their getting so many at once. Dollars are here now as plenty as dirt—

John Watts wrote Johnson from New York City in January of 1769 (Johnson 1928, 6: 587), “The Proprietary Agents were put to it to raise so much Silver, I was lucky enough to collect near One half of it by Bill on philad. when Nothing else wou’d do it.” Finally, William Gamble wrote Johnson from Albany, New York in December of 1772 to transmit an account of monies received (Johnson 1933, 8: 675-677). Out of the 910 coins listed in the account, 94 percent were silver coins and 6 percent were gold coins. In terms of value, 68 percent of the coinage value was in silver and 32 percent was in gold.

And lest the reader think that the Papers of Sir William Johnson (Johnson 1921-1965) provide the only anecdotal evidence in support of Grubb’s estimates (2004, 340) and against Michener and Wright’s claim (2006c, 12), let me offer a totally unrelated anecdotal quote from this period. Johann Carl Büttner was a German indentured servant in the Delaware Valley from 1773 to 1776. In his memoirs he wrote, “At that time there was current in Pennsylvania a good deal of paper money. . . . Also there were whole and half coins of Spanish silver money, which were chopped into four and sometimes into eight pieces. These pieces were used as change and business currency.” (Klepp, Grubb, and Pfäelzer de Ortiz 2006, 230) Büttner never mentions any gold coins being around. Büttner’s characterization of the composition of the money supply in Pennsylvania in this period is consistent with the relative quantitative estimates in Grubb (2004, 340) and directly at odds with what Michener and Wright claim (2006c, 12).

The Papers of Sir William Johnson (Johnson 1921-1965) also provide evidence to overturn Michener and Wright’s claim that the paper money of different colonies circulated freely as currency among the general populace of neighboring colonies. Benning Wentworth wrote to Johnson on June 17,
1755 from Portsmouth, New Hampshire regarding participation in the combined colonial assault on Crown Point, “The money Granted for Carrying on the Expedition being paper, it has no Currency either in the Massachusetts or in New York Governments, for which reason I have been obliged to March the regiment by Land to Connecticut River, from thence to March over to Crown point to Joyn the main body of the Troops under your Command” (Johnson 1921, 1: 606-607). As such, the paper money of one colony did not circulate freely as currency among the general populace of neighboring colonies.

Now all this evidence is anecdotal, and I am the first to put little definitive weight on anecdotal evidence, though there is a fair amount here. On the other hand, Michener and Wright live and die by anecdotal evidence. So by their own methodological standards the evidence here utterly demolishes their position and strongly supports Grubb’s quantitative estimates of the relative composition of the money supply in the middle colonies (Grubb 2004, 340-341).

**RHETORICAL GAMES (II)—THE “APPENDIX” GAME**

Michener and Wright (2006c, 253-271) relegate the bulk of their rejoinder to an appendix (18 out of 20 pages of text). As justification for this behavior Michener and Wright (2006c, 252-253) say, “In the lengthy appendix here, we address the remainder of Grubb’s reply. We relegate the details to an appendix because we believe…the tone of Grubb’s reply is the best evidence of the underlying weakness of his propositions.” Again this is an example of their preeminent rhetorical trick—preemptively accusing your opponent falsely of what you yourself are doing to insulate yourself from criticism of doing exactly that. If the reader goes back and does a head-to-head comparison on “tone” between Michener and Wright’s attacks on Grubb and then Grubb’s replies, beginning with Michener and Wright’s first paper attacking Grubb circulated on EH.NET called “Grubbing Against the Dollar,” it is clear who wins the unscholarly tone competition.

In truth, their “Appendix Game” is really just a rhetorical trick to cover up the poverty of the rest of their rejoinder where they do little to address Grubb’s replies. They just repeat what they have published before without meaningful new analysis or serious engagement with Grubb’s
arguments and evidence (the repetition ad nauseam rhetorical trick). Their opening accusation that Grubb is evasive is made to cover up the fact that their own appendix is just one big evasion.

On the issue of legal tender laws and quit rents (Michener and Wright 2006c, 253, 257-259), the assessments of Benjamin Franklin, the Pennsylvania Assembly, and the British Board of Trade all show that Michener and Wright are dead wrong (Labaree, Willcox, and Oberg 1959-1998, 6: 516-531, 9:131-173; 14: 34-39, 77-87).

On the issue of the amount of specie present at mid-century (Michener and Wright 2006c, 259-260), Michener and Wright persist in confusing specie passing through as a result of the multi-lateral balance of trade with specie retained for use by the general populace as a media of exchange for their internal commerce within the colony.

Michener and Wright (2006a, 7-8; 2006c, 253) persist in their bait and switch tactic of maintaining that evidence from a world where there is no physical representation of a monetary unit of account must have the exact same interpretation as evidence from a world where there is a physical representation of the monetary unit of account, namely pre- versus post-1723 for Pennsylvania (see also footnotes 4 and 6 above).

On the circulation of Pennsylvania pounds in Maryland, Michener and Wright (2006c, 261-262) persist in relying on a few quotations from the merchants Callister and Fitzhugh that they found in secondary sources—trying to deduce what was going on from this thin offering. They have not gone to the archives and read these sources in their entirety. They have no idea what the complete account books really look like or what the entire correspondence really says (Grubb 2006a, 63-64).

14 This is also part of their “burden-of-proof” rhetorical game. They assert that all the burden of proof falls on Grubb and none on them. Then they toss out endless ersatz economic ideas and erroneous unparsed anecdotal quotes, each of which takes up little space. Then they expect their opponent to disprove every single one. And if their opponent does not, then their opponent's position must be completely rejected. But serious scholarly analysis takes far more space (as well as research time and effort) than their cavalierly offered objections. Witness their four sentence presentation of the Tilghman quote (Michener and Wright 2006c, 262), and the amount of space and scholarly research needed above to debunk it, and debunk it so easily and definitively that any serious scholar would not have used or presented it as Michener and Wright did in the first place. They set up this rhetorical game so that no one can ever win a debate with them. Now if we reject their burden-of-proof game, and level the playing field, then Grubb has now shown so many examples where their objections are not only erroneous but fallacious that it may serve as a blanket rejection of all their objections, at least until they take on the burden of proof themselves of actually doing serious scholarly research on these issues. See also Michener and Wright (2005); Grubb (2005, 2006a, 48).
Michener and Wright (2006c, 266, 268) persist in repeating their estimate of monetary composition based on probate records. I challenge any reader to figure out how they got their numbers. Once you do and see their machinations, you will know.

Michener and Wright (2006c, 267) persist in making excuses for their misuse of, and then persist in misusing, Mazzei (Marchione 1983, 1:325-326). All I can say is go read Mazzei (Marchione 1983) and then read how Michener (1988) and Michener and Wright (2006a, 2006c) use and characterize that material. Then you will know.

Finally, this is Michener and Wright’s third publication (2005, 2006a, 2006c, 254-255) attacking a yet-to-be-published paper of mine (Grubb 2006b). I would hazard to say that this behavior is unprecedented in the scholarly journals. When the reader finally has that paper available to read they will be surprised by what is presented there and by Michener and Wright’s less-than-forthright characterization of it.

CONCLUSION

This debate is about more than the nature of money in colonial America. It is about the nature of scholarly discourse (Grubb 2001). Is scientific progress in economics fostered by debate over alternative models and their empirical testing in an open forum, or is scientific progress a myth and the goal is to monopolize society with your truth? Or put another way, there are two types of Chicago economists—those who believe in the marketplace of ideas and those who believe in forestalling others from the market and engrossing it for themselves. Which is to prevail, and what are the welfare consequences?

I see competition in the marketplace of ideas as the most important engine of scientific progress and the principal means to ensure that scholarship has some semblance of integrity. The latter position—often frequented by dictators, ideologues, and religious zealots—tempts their adherents into unbounded behaviors too easily justified and excused in their minds because they know that they possess the truth.
REFERENCES


ABOUT THE AUTHOR

Farley Grubb is Professor of Economics and NBER Research Associate, University of Delaware, Newark, DE. Link to online vita; e-mail: grubbfr@lerner.udel.edu.
Abstract

WHEN SWEDEN WAS HIT BY A DEEP ECONOMIC CRISIS IN THE fall of 1992 and was forced to abandon the pegged exchange rate, the government appointed a commission of primarily university economists under the chairmanship of Assar Lindbeck, the most well-known economist in the country at the time. The commission was assigned to present an analysis of the problems facing Sweden. Appearing in March 1993, the report had a major impact in the media and in public debate. Several of its 113 proposals eventually served as a source of inspiration for political action.¹

¹ An English translation of the commission report is available in Lindbeck et al (1994).

* Department of Economic History, Lund University
** DG ECFIN, European Commission, Brussels
This paper has evolved from a book chapter published in Swedish as "Hur såg de stora nationalekonomerna på sin roll i samhällsdebatten?", chapter 4 in Jonung (1996). It is an updated and revised version of Chapter 3 in Bellet et al (2005). Geoffrey French made the translation from Swedish (including quotations from Swedish language sources). We have benefitted significantly from the views of Michael D. Bordo, Bruno Frey, Torsten Gårdlund, Rolf Henriksson, Christina Jonung, David Laidler and Bo Sandelin. We owe a special debt to Hans Brems and Mats Lundahl for their insightful comments.
This event illustrates the strong position of the economics profession in Swedish society—no historians, philosophers, management consultants or former political leaders were considered for this task, as may have been the case in other countries. Economists play a prominent role in public debate in Sweden, many appear on radio and television, write for the daily press, magazines and books, and serve as experts on government inquiries and commissions. In Sweden, economists probably have more influence than any other category of social scientists. In other countries there is usually a wider gulf between academically active economists and the world of politics and the media, more so in the United States than in Europe.

In Sweden, the high standing of the economics profession has existed for a long time. A significant number of Swedish university economists in the twentieth century have been influential participants in the public exchange of opinions. Some have made the move from policy debate into practical politics. The founders of economics as a scholarly discipline at Swedish universities, first of all Knut Wicksell, Gustav Cassel, and Eli Heckscher, addressed the general public as popular educators and debaters.

Here we concentrate on five outstanding—and internationally renowned—Swedish professors of economics: Knut Wicksell (1851-1926), Gustav Cassel (1866-1945), Eli Heckscher (1879-1952), Bertil Ohlin (1899-1979) and Gunnar Myrdal (1898-1987)—with a view to ascertaining how they regarded their own and their profession’s role in public debate.

These five economists hold a central position in the development of economics in Sweden in the twentieth century, laying the foundations for modern Swedish economics. They represent two generations. The first generation, of Wicksell, Cassel, and Heckscher, with a clear set of values and eager to promote their views and policy recommendations. The second generation, Ohlin and Myrdal, with equally clear values, and critical, at least in Myrdal’s case, of the older generation’s, i.e. their teachers’, lack of clarity in defining the boundary between scholarship (science) and politics, yet

---

2 One confirmation of this is provided by a postal inquiry addressed to all Swedish professors of economics in 1989. The aim of the survey was to chart their contributions to the daily press, to the journal Ekonomisk Debatt, to SOU (Statens Offentliga Utredningar - Swedish Government Reports) and to similar activities. The postal replies disclosed an extensive external activity according to Jonung (1992: 42). See also Sandelin (2000) on the public activities of Swedish economists in the post-War II period.

3 The word economist refers henceforth to economists employed at universities.

4 We exclude several prominent economists such as David Davidson and Erik Lindahl, active roughly at the same time as Wicksell, Cassel, Heckscher, Ohlin and Myrdal, because they were not as visible in public debate.
most willing to move from the world of scholarship into public affairs and politics. A major line of demarcation between older and younger is marked by Myrdal’s book from 1930, *Vetenskap och politik i nationalekonomien* (*The Political Element in the Development of Economic Theory*).

The five represent a broad political spectrum. Taking “liberal” in its European meaning, Cassel was a right-wing liberal with a strong conservative leaning; Heckscher an ultra-liberal deeply influenced by the great British nineteenth century economists, and Wicksell a radical liberal. Each held back from direct party-political commitment. Ohlin was a liberal, classifying himself as a “social liberal” to distinguish him from classical or old liberalism, and Myrdal was a social democrat. Both engaged in party politics; Ohlin as the head of the Liberal Party (1944-67), and Myrdal as Member of Parliament and Cabinet Member (1945-47) for the Social Democratic party.

The five were deeply engaged in public debate. They produced a copious stream of books and articles addressed to colleagues, politicians and the public at large, gave lectures and participated in debates. One measure of this high ambition is the number of articles in the daily press: Wicksell published about 450 articles, Cassel about 1,500 in *Svenska Dagbladet* alone, Heckscher about 300 articles in *Dagens Nyheter* alone, Ohlin about 2,000 articles, chiefly in *Stockholms-Tidningen* (a large number of them before he moved into party politics), and Myrdal about fifty articles. Between them, the five economists published a round total of 4,300 articles in Swedish daily newspapers, mainly during the first three decades of the twentieth century. This avalanche of words coincided with the establishing of political economy as an academic discipline with a powerful influence on public opinion and politics.

The five had an important platform in the meetings of the Swedish Economic Association to spread their views to the public. Here they confronted politicians, policy-makers, businessmen, and fellow economists (Henriksson 2001). The proceedings were covered in the press and later published. (Table 1 summarizes some basic information on the quintet.)

---

5 *Svenska Dagbladet*, *Dagens Nyheter* and *Stockholms-Tidningen* were daily newspapers with wide circulation, published in Stockholm. Data on the number of articles are from Jonung (1991).

Table 1: Profiles of Knut Wicksell, Gustav Cassel, Eli Heckscher, Bertil Ohlin and Gunnar Myrdal

<table>
<thead>
<tr>
<th>Name</th>
<th>Main academic post</th>
<th>Major scholarly works</th>
<th>Newspaper articles</th>
<th>Political outlook</th>
<th>Role in party politics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knut Wicksell (1851-1926)</td>
<td>Lund University 1901-1916</td>
<td><em>Value, Capital, and Rent, Interest and Prices; Lectures on Political Economy</em></td>
<td><em>ca 450</em></td>
<td>Radical liberal</td>
<td>None</td>
</tr>
<tr>
<td>Gustav Cassel (1866-1945)</td>
<td>Stockholm University 1903-1936</td>
<td><em>Nature and Necessity of Interest, Theory of Social Economy</em></td>
<td><em>More than 1,500</em></td>
<td>Right-wing liberal</td>
<td>None</td>
</tr>
<tr>
<td>Eli Heckscher (1879-1952)</td>
<td>Stockholm School of Economics 1909-1945</td>
<td><em>The Continental System, Mercantilism, Economic History of Sweden</em></td>
<td><em>More than 300</em></td>
<td>Classical liberal</td>
<td>None</td>
</tr>
</tbody>
</table>
The appearance of these five economists on the public arena invites a number of questions such as: Can an economist working in academia be scientifically objective and politically committed at the same time? Ought he to confine himself to scholarly discussion or attempt to play a pedagogical public role as well? Ought the economist to strive for expertise within a narrow field or for broad knowledge and understanding?

We seek to chart how the five economists themselves answered these questions concerning the role of the economist in public debate. We seek their answers first in their own writings, chiefly “mature” works and autobiographical articles or memoirs, and, second, in biographies and other writings about them. In our conclusions we consider briefly the present tendency of university economists of shunning away from public debate.⁷

---

**KNUT WICKSELL—EDUCATOR OF THE SWEDISH PEOPLE**

Knut Wicksell is considered Sweden’s foremost economic theorist of all time. He was a great innovator in economic theory, principally monetary theory, public finance, and capital theory. But he also devoted himself to economic, social and political issues on a broad front. Wicksell stood out as a social critic of rank, questioning established institutions such as marriage, the church, the monarchy, and the military. All his life he enjoyed a well-founded reputation as an independent radical, always ready to defend and advocate views that were regarded as extreme in public debate.

Some have found it perplexing that Wicksell could simultaneously play these two roles: on the one hand, the unassuming academic, on the other, the vociferous agitator. There are two different interpretations. The first, which we can call the Gårdlund-Ohlin line, sees Wicksell as having a dual nature. The second, the Lindahl-Åkerman line, points to Wicksell’s passionate commitment to everything he undertook.

Torsten Gårdlund (1956, 371) formulates Wicksell’s dual nature in these words: “Knut Wicksell’s character sometimes gives the impression of a strong inner conflict, almost a split personality in fact. He was a wild

---

⁷ The great five were all men. Women entered Swedish economics at a much later stage in the 1970s. For an overview of women’s entry into economics and participation in economic debate see Jonung and Ståhlberg (2002) and (2003).
agitator and an objective scholar at one and the same time.” Bertil Ohlin (1972, 558) muses along the same lines:

To me it is a riddle that Knut Wicksell, who for most of his life was a fanatical representative of extreme opinions in the social debate, could present a completely different personality in the scholarly context. During the period when I knew him he was the diffident seeker after scientific truth.

In a more detailed discussion of Wicksell’s nature, Gårdlund (1958, 200) writes:

In the political sphere he recognized no authority. No institution was sacred; no hallowed law could prevent him applying the test of democracy and utility. To plead tradition in politics was ‘obscurantism’, to suppress criticism was ‘under any conditions an evil’. But as a scientist he submitted himself voluntarily and with remarkable consistency to the traditional demands of the scientific world. His ideal scientist was one who sought the truth sincerely and without prejudice and who presented his findings objectively and modestly.

Erik Lindahl (1953, 304-305)—supported by Johan Åkerman (1956)—preferred to emphasize that Wicksell’s scholarly and political activities were founded on the same passionate commitment.

Wicksell was both a scholar and a social reformer, the latter on the ideological plane in the capacity of popular orator, debater and author of controversial pamphlets on social policy. One feels that it should have been an abrupt reversal for him to tear himself away from his hyper-theoretical work at the desk in order to make an agitator’s speech to a demonstration meeting or to speak from the platform at a young socialists’ meeting. But Wicksell performed the one task just as intensively as the other. ... And in his case it is evident that the one activity had as fruitful an effect as the other. [It was] Wicksell’s social interest which impelled him to take up economics. . . . On
the other hand, his social preaching achieved greater
coherence through being buttressed by scholarly arguments.

Mats Lundahl (2005b, 170-171) has recently added to the discussion
by supporting the view that there was basically only one Wicksell. In a
pathbreaking interpretation of Wicksell’s analysis of poverty and population
growth, he argues:

The result of putting all the bits and pieces of Wicksell’s
scattered analysis of population growth together is
astonishing. Far from confirming the conventional wisdom
that what he wrote on the population question was
mechanical and simplistic, it turns out that the exercise
results in a coherent general equilibrium framework which
very much resembles the specific factor models of
international trade developed by Ronald Jones (and Paul
Samuelson) in 1971. Within this setting Wicksell handled
factor growth (population, natural resources, capital),
technological progress, tariffs and factor movements. In
this, he stands out as a precursor of the modern theory of
international trade.

Concluding his book on Wicksell on poverty, where this argument is
developed in depth, Lundahl (2005a, 102) states:

All the time, however, what Wicksell did in the public
sphere had a solid foundation in his scientific thinking, and
despite what has been alleged, this is the case also with his
views on population growth and poverty.

Lundahl reveals the close connection between Wicksell’s popular
writings and his scientific work. This suggests that the two interpretations
on Wicksell’s nature are not contradictory. It is partly a matter of where the
emphasis is placed: on the unifying commitment of penetrating deep into
social and economic issues or on the dissimilar modes of expression in
public debate and in scientific work.

One explanation for the view that Wicksell had a dual nature can be
sought in the fact that he had a natural bent for abstract and logical
thought—he was the archetype of the deductively-working scholar. “The
abstract intelligence of which he early showed proof must have helped
greatly to make him a man of principle and opinion,” Gårdlund (1956, 364) contends. In all likelihood, Wicksell’s acting in accordance with his nature produced differing effects on his scholarship and his politics, respectively. To argue with uncompromising logic is one thing in an academic discipline where the dominant tradition prizes logic above adherence to reality, but it is something else in a political reality, which builds on historical circumstances rather than on logical designs. Wicksell’s critics often argued that in his economic policy recommendations he tended to think too straightforwardly, without taking account of the complications entailed in translating ideas into practice.

Wicksell’s close friend, Hjalmar Öhrvall, put his finger on this sensitive spot in one of his many letters to Wicksell.

I can understand quite easily that a man can have a definite, contrary idea about a scientific proposition and be willing to defend it against the whole world; but how it can be possible to be as definite on political, strategic or similar questions, I do not understand. (Quoted in Gårdlund 1958, 242)

On top of this there was the fact that in his political appearances Wicksell was driven by his urge to provoke scandal. Actually, he used the scandal as a method of bringing out his message to the public.

Wicksell might be dubbed a radical liberal. He fought for a more equal distribution of wealth and income. But he did not share the Marxist notion of capitalism as a system of exploitation. He believed that on the whole the market economy would give the best economic outcome and that the distribution problem could be solved by means of education and inheritance taxes. He cherished a deep sympathy for the social democratic cause and at one time in the 1880s nearly became a member of the party, but he preferred to function as “a radical taskmaster untrammeled by party ties” according to Gårdlund (1958, 307).8

---

8 According to Erlander (1972, 122), Wicksell was invited to speak at the twenty-fifth anniversary of the local labour party organization in Lund in 1926. Wicksell expressed his appreciation of being the keynote speaker but added: “I do not belong to the herd. I am a sheep all by myself.”
One of Wicksell’s clearest manifestos is to be found in his inaugural lecture as professor in Lund “Ends and Means in Political Economy.”\(^9\) Here Wicksell (1904, 460, 470) discusses why it is so difficult to agree on anything in economic science. Is it because the problems of political economy are so difficult that we have not found the right way to solve them? Or

is it perhaps the various scholars’ individual sympathies and antipathies, their diverse political ideals, their conception, in a word, of the aim of practical social and economic activity which lends its color even to their treatment of theoretical questions?

Wicksell moves quickly to the heart of the question:

That the aim of economic activity must be the greatest possible prosperity of society, individually and collectively, on that point we are all formally agreed; but what is meant here by society? Shall our endeavors embrace all classes, races, linguistic groups, creeds, nations?

Wicksell contends that the answer is yes, and he gives expression to his utilitarian view that “our aim here on earth is to extend the greatest possible happiness to all” (470). If economists would only adopt this view, they could also attain unity.

Wicksell believed that modern economics had passed the short-sighted vision of the harmony economists and was no longer the creature of any special interests.

I have always regarded it as a criterion indicating that modern theoretical economics is on the right road that the scorn for everything that economists stand for, so common among the working class in bygone days, has died away . . . the workers probably sense instinctively that the watchword of economics has again become the unconditional quest for truth. (472)

It is clear from his writings that Wicksell viewed himself as an educator of the general public on any issue he found of interest to bring to the public. Through his many articles, lectures and comments, he wanted to foster public knowledge of the science of economics or, as he sometimes wrote, of the laws of economics, in particular of Malthus’ population theory. He was convinced of his mission as a public educator (folkbildare). His view is revealed in a letter to Hjalmar Öhrvall at Christmas 1916, shortly after Wicksell had retired from his chair in Lund and returned to Stockholm. Here, he contemplates his future as a professor emeritus.

Yes, it feels somewhat strange to have been ‘removed from office’, but not very much since, in particular in later years, I took very little part in academic life. H. Hildebrand, my old friend from school, who retired already last summer, says that he finds it so strange to wake up in the morning and ‘not have any duties’. As for me, I still find it my foremost duty to educate the Swedish people, and I must still try to do this, as long as I have the

---

10 During the controversies surrounding the chair in Lund, the critics of Wicksell opposed his propensity to make his views known to the public. Gårdlund (1956, 244) quotes a comment against Wicksell, based on his “detrimental activities against his native country as a public educator.”
strength. (Jonung, Heldlund-Nyström and Jonung 2001, 19)

Here, Wicksell puts his mission in life on paper: “my foremost duty to educate the Swedish people”.11

Wicksell was very outward-looking all his life. He wrote pamphlets, lampoons and articles, and he went out on tours giving, as he said, ‘lectures to the peasants’—and he also held the view that his intuition led him correctly in his communication with the public:

As for myself, I will say one thing in my favor; that I have a nose for what can be done. In other words, I feel that if something is obvious enough to penetrate my simple understanding, it cannot be long before it will conquer the masses—and I have never been wrong about this, although sometimes things have moved a little more slowly than I expected. (Quoted in Gårdlund 1958, 307)

GUSTAV CASSEL—THE VOICE OF REASON

In his day—during the 1920s at least—Gustav Cassel was one of the world’s most renowned economists, a theoretical innovator as well as a leading expert on current monetary problems. In the late 1920s he was introduced to the American House of Representatives as the world’s most famous economist; his only competitor for that “title” was John Maynard Keynes. According to Magnusson (1991, 134), his reputation did not rest so much on his role as innovator and theorist: “His strength lay rather in economic pedagogy, in teaching, and as an inexhaustible promoter of the subject of economics to the lay public and political society.”

Cassel started out around the turn of the century 1900 as a fairly radical liberal—he was even accused of being a socialist! However, by the end of World War I he had developed into a right-wing liberal who fought relentlessly against socialism, state enterprises, economic planning, government regulation

11 The quote is from Gårdlund (1956: 337). It is adopted as the title for the publication of about eighty previously unpublished manuscripts by Wicksell. See Jonung et al. (2001).
and the like. For him, the question of whether the scholar ought to become involved in public debate was easily answered. The scholar represented the clear voice of reason in a world governed by superstition and dilettantism. The task of the scholar of economics was to look at the whole, to elucidate the inexorable economic laws from a standpoint high above the clash of interests and thereby help the public, politicians and businessmen to see beyond their own narrow interests and time-horizons. Out there in the economic and political jungle all kinds of delusions flourished, and it was the task of the scholar to clear up the miserable mess with the shining weapon of reason.

In my scholarly work I have chiefly been seeking clarity. I have called for rational action in the life of society, with the accent both on a reason, which does not allow itself to be tied down by arid dogma or sterile party formulae, and on a robust will that is prepared to act at the right moment. These main aspects of my work have been linked together inasmuch as my efforts to bring clarity to the elements of economic science have become of crucial importance to my practical standpoint, and also because my scholarly work has been stimulated by the immense problems with which life has confronted me.

Thus writes Cassel in the preface to his memoirs *I förnuftets tjänst* (*In the Service of Reason*) (1940, 7). In his final words (1941, 455-56) he returns to this theme.

It fell to my lot to work in the service of reason. And this is a duty. It requires fidelity all through life. . . . To serve reason certainly requires humility in face of the task. But this humility is essentially different from the obliteration of one’s own personality, when one joins a party or takes refuge in the bombproof shelter of collective programs. He who fights for reason must give himself to the struggle, sticking it out though he finds himself standing alone.

Cassel wrote his memoirs after the depression of the 1930s and at the beginning of the Second World War. By then he thought that most of what he had struggled to achieve was reduced to ruins.
Everything I have fought for now lies in dust and ashes. Scholarly enlightenment counts for nothing in an age, which feels more comfortable with drivel. Free speech is suppressed and the technical means of reaching listeners are cut off or monopolized. International communications are impeded and mankind is excluded from mutual discussion of its vital affairs. . . . Sensible economic management by government is regarded as an obsolete prejudice and huge budgets are run up with nothing on the revenue side. Law is thrust aside by administrative arbitrariness, and democratic essentials are either shattered or diluted to empty formulae. Of the freedom of the individual scarcely more than the name is left, and the value of his personality is set at naught. Wherever one looks, only destruction! Destruction at least of that which I had wished to build. (Cassel 1941, 456-457)

In Cassel’s description of this heap of ruins we discern the foundations of the proud edifice for which he had fought and which bears the characteristic marks of liberal progressive optimism: individualism, free enterprise, free trade, free pricing, policies devoted to the creation of law, order and a stable currency, peace and progress (that is to say economic growth).
In other words, Cassel saw himself as a scholarly interpreter of economic conditions and relationships that were virtually to be regarded as laws of nature. He summarized the chief tasks of economic science in the introduction to his *Teoretisk socialekonomi* (*Theory of Social Economy*) (1934) under five points: (1) “to view society as a whole entity and to try to trace the total nexus of causality within this society” (social economic thinking instead of private); (2) “to bring real phenomena out into the light of day” (for example, the utilities that in everyday terms are expressed as a sum of money); (3) “to try to penetrate to the essentials of economic phenomena and situations”; (4) “to penetrate to the necessary facts of economic life” (for example the proposition that capital formation is necessary for progress); and (5) “to give a simplified picture of economic life”, in other words to formulate a theory.\(^{12}\)

In spite of his lively interest in the issues and realities of the day, Cassel, whose early training was in mathematics, was at heart a logician. The scientist in economics must, in turns, make use of deduction and induction, but the point of departure must be deductive: “The main thrust of the simplification procedure consists in isolating the essentials from the very start and reproducing them in a logically coherent presentation” (Cassel 1934, 20). One may think that, as a scholar who expressed himself decidedly, Cassel here exposed himself to a considerable risk. Suppose his theory has indeed isolated essential relationships and is logically coherent, but is not the only logical model capable of being constructed around a number of essential empirical factors and relationships, what then? The risk is lessened in Cassel’s case by his known lively interest in facts, in “reality”.

At the same time, however, he was aware that his preaching rested not merely on logic, but also on value judgments. When other economists and politicians accused him of marketing his political opinions as scholarly conclusions, he admitted without equivocation that his positions rested on certain fundamental value judgments.

If my theoretical studies have contained any subjective value judgment, then this has amounted at most to a preference for freedom and progress rather than state control of the economy and distribution of such scanty prosperity as may be available for distribution at a given moment. I have wanted to make it clear that this preference is a great common interest of all parties, both in the management of the world economy and in every

\(^{12}\) This introduction is not included in the English edition of 1932.
individual nation. Such a position may be attacked, but it cannot be denominated as party-politics in the ordinary sense. (Cassel 1941, 440)

To be pushed into a party pigeonhole was, in Cassel’s eyes, one of the worst fates imaginable.

Cassel’s strong position in public debate rested in large part on his capacity to explain economic issues clearly, simply, and elegantly so that politicians and the public could keep up with him. He himself was very aware of this capacity, and we have testimony as to how hard he actually toiled over texts that seemed to flow with their own momentum. In a letter to Ohlin he declared: “Economics is in high degree a pedagogical discipline, and an economist must be in close touch with popular psychology in order to know what ought to be said at any particular moment” (Quoted in Ohlin 1972, 107).

ELI HECKSCHER—LIBERAL BEACON

Eli Heckscher was first professor of economics and later (from 1929) professor of economic history. In fact, he established economic history as an independent academic discipline in Sweden. In his younger days he was a social conservative, but after World War I he emerged, much like Cassel, as a full-fledged liberal with strong sympathy for British 19th century economic liberalism.

Heckscher’s work was marked by the tension between scholarship and politics. Rolf Henriksson (1979, 519-520) describes the matter in this way: “In Heckscher’s work as an economist the tension between the political and scholarly sides emerges clearly. In his academic work he never relinquished the politico-ideological starting point, and in his political attitudes the scholarly dimension was always present.” Heckscher himself was highly conscious of the dilemma.

For on the one hand he [the scientist or scholar] is a citizen and accordingly has the same duty as others to form a subjective, practical opinion concerning matters which in a democratic society depend on the decisions of all citizens. . . . But his practical standpoint must
necessarily contain a purely personal value judgment, which is not that of a scholar. On the other hand, however—perhaps even first of all—his duty is that of the scientist or scholar, viz. to present objective truth to the utmost of his ability, regardless of his own sympathies or antipathies. He can—and if he is conscientious must—seek to make clear to his readers and listeners where the boundary lies between the objectively valid and the subjectively evaluated; but even to make it clear to himself is a very difficult task. (Heckscher 1936, 2)

Eli Heckscher

Heckscher elaborated on this view ten years later:

In some cases . . . the conclusion has been drawn that scholars ought to keep themselves politically neutral: on this view the representatives of economic science must not pronounce on what ought to be done but confine themselves to analyzing the actual circumstances of economic
life, in other words analyze what has happened and is happening (and possibly also state what will happen under various conditions). . . . [W]ithout doubt it is usual for economics scholars also to make frequent pronouncements about what ought to happen. This is hardly to be condemned, either, since their theoretical insights and overall view may be expected to enable them to avoid many mistakes which otherwise are easily made. Like all citizens they have the right to plead their case on the problems of society and in their special field more so than others if they understand the questions better. But the last part of their submission, amounting to a specification of claim, falls beyond their scope as scholars, at least as long as they have not clarified for themselves and declared to their public the unscientific value-premises which underlie their recommendations and advice. Of course it is never easy for economists strongly interested in the problems of society thus to split their personalities, as it were, into a scholarly and a civic half. (Heckscher and Knoellinger 1945, 25-26)

Heckscher’s argument concerning the importance of scholars’ declaring their value-premises was probably influenced by Gunnar Myrdal’s demands in the same direction.

The only general advice Heckscher was willing to offer those thinking of studying economics was:

that which applies to all honest study: to use your intelligence, to be receptive to all enlightenment but not to take anything for granted beforehand, whether your own previously-formed opinions or statements made by those who are regarded, or want to be regarded, as authorities, whether in theory or practice. (Heckscher and Knoellinger 1945, 16)

Nevertheless Heckscher (1936: 4) considered that an adherence to neoclassical theory, to concepts of equilibrium and scarcity, and to an atomistic approach was a great help along the road which “enforced objectivity and respect for given scientific assumptions.” Indeed, he declared outright that economic theory as it evolved in the 1930s, i.e. the emergence of the theories and economic policy recommendations
associated with the Stockholm School and Keynes’ *General Theory of Employment, Interest and Money* (1936), had expanded the scope for “cobbling up theory so as to suit the cobbler’s own political or social viewpoint.”

However, for anyone not starting out from the same “given” (liberal) premises, it was an obvious step to suspect a subjective/political bottom below the objective/scientific surface. The economist and conservative politician Gösta Bagge reproached Heckscher for “being unwilling to regard liberal politics as politics” but presenting liberal desires as being “objective” or “economic”. The economist and left-wing politician Gunnar Myrdal likewise accused Heckscher of promoting liberal policy in the guise of objectivity (Carlson 1994, 21).

Thus, Heckscher stood on liberal ground. His uncompromising attitude and steadfastness amidst the ideological and political storms blowing in the 1920s, 1930s and 1940s made him, in Ernst Wigforss’ (1951, 155) words of reluctant admiration:

> something of a beacon when navigating the waters of economic policy, inasmuch as he was regarded as a representative of a reasoned and coherent economic ideology whose liberal argumentation the socialist could not by-pass but had to consider and pronounce upon.

In his later years Heckscher came round to the view that “everything was better in the old days.” By “in the old days” he meant the liberal era from the middle of the nineteenth century up to the outbreak of the First World War, or what Heckscher called “nineteenth century economic order.” Heckscher’s stance now was not merely a manifestation of the conservatism which easily comes creeping in with age but a quite reasoned reaction to the horrors visited on the world during the twentieth century—First World War, the Great Depression, Second World War—especially if, like him, one regards these horrors as being caused by high-handed politicians and national power-plays and not, as socialists saw the matter, by capitalist rapaciousness and market anarchy.

Like Cassel, Heckscher tended to think that the economics scholar must begin at the deductive end by abstracting and theorizing, by evolving for himself a view of “the elements of what is common to all economies.” To avoid the temptation of regarding the economic organization of his own

---

13 Ernst Wigforss was a leading socialist and social democratic ideologue. He was minister of finance 1925-26 and 1932-49.
time as a manifestation of ineluctable and universally valid laws he also ought to obtain “concrete knowledge of the external phenomena of the life of the society, and most especially of its economic life, during different eras;” in other words, he ought to study economic history (Heckscher and Knoellinger 1945, 18-24). Heckscher strove as an economist and economic historian for both breadth and historical depth in the study of economics.

Towards the end of his life Heckscher, like Cassel, felt himself to be standing in a world laid to waste; one of his late works is entitled Ödeläggelsen av 1800-talets husbäldning (The Destruction of the Nineteenth Century Economic Order). During the 1930’s, unlike Cassel, he became less active as a public debater and molder of public opinion—not because he had capitulated to the “new” currents of economic policy opinion in any way. Rather, he kept a lower profile for two reasons: first, he was concentrating on his research in economic history after obtaining a personal research professorship in economic history in 1929; second, after Hitler’s rise to power in Germany, he feared that polemical contributions by a Jewish scholar would spark anti-Semitic reactions.

When the Hitler regime collapsed, Heckscher began to play a leading role again in the economic policy debate as one of the most incisive figures among opponents of the economic planning advocated by Myrdal, Wigforss, and other leading social democrats. One of his crucial arguments in the debate was that “scholars are free because they have access to a free market, they can publish books, they can write in the newspapers, they can make their voices heard generally in the national life, because the means of production are free” (see Röster i Radio 1945, no 28, 36). State ownership or direction of the means of production would end up with standardization of opinion and intellectual dictatorship. In this respect Heckscher followed the same line as Hayek.

Heckscher (1944, 92-93) believed that economists of his own generation had had very limited opportunities of influencing economic policy directly. “As far as I can understand, the only point at which we managed to make any mark was in influencing public opinion.” But the economists of the next generation were deeply involved in party and government machinery. Perhaps they had succeeded in influencing events in certain respects, Heckscher argued, but:

I believe one may say that independence and influence stand in an inverse relationship to one another; and when

See Heckscher (1948).
influence has now become so great, then scholars more or less inevitably must feel themselves fettered by every possible consideration from which we were free.

BERTIL OHLIN—SWITCHING ROLES ELEGANTLY

Bertil Ohlin enjoyed a brilliant academic career, being professor of economics at the University of Copenhagen by his 26th year. A decade later he began an astounding political career. For 23 years he was head of the Liberal Party (folkepartiet) and an almost indefatigable contender against the Social Democrats, whom he never was able, however, to defeat in an election.

Ohlin began as a scholar, then wandered back and forth for a time across the boundary between the academic and political worlds, eventually becoming a full-time politician. Ohlin and Cassel are the only two of the five economists who wrote detailed memoirs. Ohlin might be expected to have been one of the economists with the most to say about economists’ participation in public debate and the art of tightrope-walking between scholarship and politics. As a matter of fact, he did not say much, at least not in his memoirs. As long as the discussion centers on himself as a scholar and a newcomer to politics it is always “I” saying and doing this and that; when it turns to himself as party leader, the wording changes so that it becomes “we” saying and doing things.

15 Working with his monumental study of Bertil Ohlin, Sven-Erik Larsson (1998) was unable to find Ohlin discussing in earnest the division of roles between the politician and the scholar. Larsson stated in personal communication with Lars Jonung that this had to do with “the reticence that Ohlin always displayed with respect to personal problems.” Ohlin, the family man, the professor and the politician, is portrayed in a number of contributions in Findlay et al. (2002). They provide the best picture of him available in English.
From the outset Ohlin was tuned into public debate. He published his ideas from the early days in Svensk Handelsstidning, Svenska Dagbladet and elsewhere. From 1925 to 1929 he was professor at Copenhagen. “In Danish newspapers there is a mass of evidence of his efforts as economic debater and popular educator,” writes Larsson (1998, 53). When Ohlin returned to Sweden as professor at the Stockholm School of Economics in 1930, he continued as an assiduous producer of articles on economic questions. The articles flowed from his pen in the early 1930s in a stream of swiftness hard to surpass.16

Ohlin’s transition from scholar to politician seems to have taken place simply and elegantly—and scarcely unexpectedly.17 It may well have been relatively painless to begin with: Ohlin, as an independent thinker, presented ideas about what a liberal party ought to be doing, whereupon the party approached him with the request that he put his ideas into effect on its behalf. But conflicts must still have arisen sooner or later between the independent thinker and a political party’s need for troops and tactics.

16 An example of Ohlin’s enormous productivity: in the years 1932-1943 he published on average almost 70 articles a year in Stockholms-Tidningen alone. See Carlson et al. (2000). Ohlin’s newspaper articles have so far not been collected and analysed by scholars. An attempt to convey Ohlin’s “popular message” on the Depression of the 1930s is made in Carlson and Jonung (2002).

17 Professor Thomas Thorburn recalls from his time as a student at the Stockholm School of Economics with Ohlin as his teacher in economics: “His students were guessing at this time as to whether he would join the Social Democrats or the Liberals, but expected Ohlin to go into politics.”
Ohlin declares in his memoirs that he counted himself as one of the liberal camp even in his youth. The keystone of his liberal outlook, exactly as with Cassel and Heckscher, was the conviction that an economic system based on individual property rights and competition would foster high growth, and that, in the long run, growth was crucial to the development of prosperity. One of his expressed goals was, just as with Cassel, to “counteract economic superstition” (Ohlin 1972, 61-62, 184). Gradually, he came into conflict with his teachers, especially with Heckscher. Ohlin argued that a liberal social order was compatible with an “active” countercyclical policy and, to a certain extent, with economic planning. Heckscher was strongly opposed to such a view.

In his memoirs, Ohlin (1972, 159) discusses very briefly the relationship between scholarship and politics in the course of his references to Myrdal’s work *Vetenskap och politik i nationalekonomien (The Political Element in the Development of Economic Theory)*.

It became of importance to Scandinavian economic scholarship through its urging of caution when formulating concepts and asserting objectivity. Of course, it is self-evident that an opinion about what ought to be done—that is to say on social policy—is based on subjective value judgments. Caution is advisable here too. What these value judgments are ought to be stated, and naturally terms with a definable content should be used as much as possible. Unfortunately both these requirements are often impossible to fulfill in political work, which I sometimes found embarrassing later on.

Ohlin thought that “facts could speak for themselves,” in a way that Myrdal hardly would have accepted. He eventually published a book, *Obekväma fakta (Uncomfortable Facts)*, where the ambition was to “clarify how

---

18 After a debate in Oslo in 1935, Heckscher wrote to Ohlin (in a letter of 21 June 1935) that there was a risk that the “present powers might be able to get you to go along with just about anything” and that Ohlin would not fit in with the Liberal Party but ought to follow Myrdal's example and join the Social Democrats. Ohlin (3 August 1935) replied very caustically. The relationship between Ohlin and Heckscher was never again what it had once been. Letters from and to Eli Heckscher quoted in this article are available at the Royal Library in Stockholm.
things really are” in order to stimulate a more constructive public debate by educating ignorant debaters about important facts, and by forcing debaters who are aware of these facts but find them uncomfortable to take them into consideration (Ohlin 1971, 7-10).

Ohlin felt a calling to serve the public, which may perhaps explain why he donned the political mantle instead of continuing to develop as a theoretician, even though theoreticians also serve the public if their ideas can be translated into practical applications or if they can expunge ideas which ought not to be put into practice. “One problem which I addressed early on,” Ohlin (1972, 97) writes, “was whether one ought to strive primarily to be useful, to perform services for the public by one’s work. Or ought one to seek to develop oneself?” He chose the former alternative.

That Ohlin was animated by a desire to serve the public, and that he considered himself a strong enough character not to have to compromise his own (scientifically based) views, is apparent from one of his letters (dated 3 August 1935) to Heckscher.

For my own part I only want to add that if I ever go into politics more actively, it will be for two reasons. Firstly, to try to help to improve the situation of the poorest people in the country, which the Liberal Party ought to be able to do better than the trade union-linked Social Democratic Party. Secondly, because our political life needs upright and independent persons, which is what I should try to be.

GUNNAR MYRDAL—SOCIAL DEMOCRATIC WONDERBOY

Gunnar Myrdal was an economist who held views on most subjects and delighted in sharing them with the world; in this respect he followed in the footsteps of his teacher Gustav Cassel. His views concerned fiscal and monetary issues, economic planning, commercial, social and housing policy, population, education and development questions, constitutional issues, etc. In fact, he considered it his duty as a scholar to have opinions about most things. “As an institutional social scientist, I believe in principle that everything can be explored, and also that everything which is important ought to be discussed and explored” (Myrdal 1982, 138).
Myrdal believed that economists could develop their capacity for questioning and searching for truth only by attacking on many fronts, acquiring a wide-ranging knowledge of society and taking part in public debate. Economists who isolated themselves “in the little model world of their own and their colleagues” risked becoming narrow-minded, uncritical and conservative scholars. In former times economists had begun their active lives as businessmen, mathematicians, historians, philosophers or lawyers. “Their changing over to economics was regularly caused directly by a strong involvement in the problems of society. Economists almost never started off as economists.” When economists began life as economists, they increased the risk of becoming narrow specialists who never questioned the fundamental value judgments and traditions of their own disciplines (Myrdal 1973, 71).

The question of deduction and induction was closely associated with these issues. Myrdal held that it was necessary, on the one hand, to think in abstract terms and to find the essential elements of existence by the deductive route, but, on the other hand, to keep in close contact with reality, that is, to work inductively.

The further away a scholarly opinion is from direct observation and the more abstract and ‘theoretical’ it is, the
more defenseless it becomes against insidious opportunist errors of judgment. In economics, model thinking in particular creates scope for systematic biases. . . . But of course all social studies must nevertheless aim at generalization. It is thus important to be able to think concretely at the same time, as I learnt from Gustav Cassel. (Myrdal 1982, 265)

Myrdal had a clear party-political profile. In his youth he belonged to the conservative camp, but turned in the early 1930s into an active Social Democrat and made a rapid career: adviser to the Minister of Finance, Member of Parliament and Minister of Trade for two years immediately after World War II and author of various policy reports. Thus, he went further than his mentor and friend Cassel, who confessed that he had a general liberal outlook but declined to accommodate himself to any party line.

Myrdal’s solution to the scholarship-politics dilemma, as is well known and highlighted for example by Klein (2006), was for the scholar to work on the basis of explicit value-premises. He began this line of thought in his *Vetenskap och politik i nationalekonomien* from 1930, a bold and challenging book published when he was just 31 years old. In the preface to the English edition of this work Myrdal (1953, vi) recalled why he wrote it.

The setting was the Swedish economic discussion in the late “twenties”. . . . Particularly after Knut Wicksell’s death in 1926, a very uncompromising laissez-faire doctrine dominated the teaching of economics in Sweden. Gustav Cassel, Eli F. Heckscher and other Swedish economists who are less well known abroad . . . were prolific writers. They also wrote popular books and they contributed to the daily press. They had an enormous influence, in Cassel’s case far beyond Sweden. This book was planned as a frontal attack on the dogmas of the older generation and it was originally meant to be a popular exposition.

---

19 So far it remains an unresolved issue why, around the age of 30, Myrdal left his conservative outlook and became a Social Democrat. In 1932, the same year as the Social Democrats entered government, he and his wife Alva became party members. In personal communication with Lars Jonung, Myrdal stated that his move was influenced by Alva’s studies in psychology.
The book from 1930 was thus mainly devoted to criticism. Myrdal wanted to cut away all “metaphysical elements.” This ambition was, paradoxically enough, both inspired by and directed against Cassel—inspired by Cassel’s attack on value theories and directed against Cassel’s laissez faire “prejudices.”

Myrdal propounded the view that the scholar should state his value judgments and would then be able to reason objectively, scientifically, logically, and rationally. Little by little he came round to the opinion that this book’s conception of the existence of some solid and ‘objective’ body of economic theory was mistaken and that value-premises are needed even to establish facts, not merely for drawing political conclusions. The value-premises selected have an influence on how research is approached, in other words on what is studied and what questions are asked, on the focus and conduct of analysis and the conclusions drawn (Myrdal 1973, 11 and 1982, 265).

Myrdal had planned The Political Element as “a frontal attack on the dogmas of the older generation” (Myrdal 1953, vi). He expressed his criticism of the older generation without reserve in a letter (dated 26 December 1934) to Heckscher:

As regards the problem of science and politics I shall confine myself to the following: Here for a generation you have pursued a liberal policy in objective guise, that is to say without declaring the moral and political value-premises: ‘objective’ discussion, in your opinion. When we are writing on policy we say so, and also attempt to state our value-premises, and then you talk about fanaticism. You have evidently not understood our critique of value. You are still stuck in your ‘objective’ policy, which shuns moral judgments.

20 For a more comprehensive exposition of Myrdal’s view of his Swedish teachers and colleagues, see the postscript to Myrdal (1958).

21 Heckscher wondered (in a letter of 18 December 1934) whether Myrdal “does not at bottom lack the temperament of a scholar and so ought to make his main occupation that of an agitator.” It was precisely in his critique of value judgments that the subjectivism of Myrdal’s scholarly work revealed itself. “Why does Vetenskap och politik i nationalekonomien concentrate its fire solely on liberal ideology and not on socialist also?”
As time went by Myrdal came to the conclusion that it was impossible to proceed by establishing objective scientific knowledge, adding values, and drawing political conclusions.

This implicit belief in the existence of a body of scientific knowledge acquired independently of all valuations is, as I now see it, naïve empiricism. Facts do not organize themselves into concepts and theories just by being looked at; indeed, except within the framework of concepts and theories, there are no scientific facts but only chaos. There is an inescapable a priori element in all scientific work. Questions must be asked before answers can be given. The questions are an expression of our interest in the world, they are at bottom valuations. Valuations are thus necessarily involved already at the stage when we observe facts and carry on theoretical analysis, and not only at the stage when we draw political inferences from facts and valuations. I have therefore arrived at the belief in the necessity of working always, from the beginning to the end, with explicit value premises. (Myrdal 1953, vii-viii).  

Myrdal was inspired by Cassel’s endeavors to dispose of utility theory. The efforts made by both Cassel and the Uppsala philosopher Axel Hägerström to avoid “metaphysical speculation” encouraged Myrdal to start grappling with the problem of science and politics. His final solution to the problem was a pessimistic one regarding the possibility of conducting value-free, objective research. Value judgments permeate everything. He stubbornly insisted on calling economics by its old name of political economy.

Over and above that of subjective value judgments, party-political activity presents another problem: that is the problem of being forced to adapt oneself to the value judgments of others. Myrdal apparently believed that when he joined the Social Democratic Party it was so open and tolerant that he would never need to adapt himself to any party line. “With the open-minded character which the party then had, I never had to feel myself subject to the kind of party loyalty which could restrict my freedom of opinion in the slightest degree” (Myrdal 1982, 224). Nevertheless he was

---

22 For a discussion of Myrdal and his approach to values in social sciences see, for example, Peltier (1992) and Wundrak (1991).
aware of the dilemma, which became acute in 1945, if not before, when he became Minister of Trade and a cog in the machinery of government.

When I agreed to become a member of the new entirely Social Democratic government, I remember I was profoundly gripped by the feeling that it must signify a change in the conditions of my work. From having been a wholly independent scholar active in the social and economic field I now became a member of a collective decision-making group. . . . I had to set a boundary to what I personally thought and even to what I could reveal publicly concerning any possibly divergent opinion I might have. I was strongly sensitive to the limitation on what I could declare openly as my own opinion or divulge about what had preceded government decisions. I had never before needed to observe any such limitation. (Myrdal 1982, 225, 229)

In his ambition to speak in a simple way to the people, Myrdal did not lag behind his teacher Cassel. In his book *Objectivity in Social Research* (1969, 41-42), he expressed his concern over recent trends.

The great tradition in social science and, particularly, in economics has been for the social scientists to take a direct as well as an indirect responsibility for popular education. There is a recent trend, with which I must register my dissatisfaction, to abandon this great tradition. Through generations even the greatest scholars—and they especially—managed to spare time from their scientific work to speak to the people in simple terms that laymen could understand. Yet too many social scientists today are increasingly addressing only each other. This trend to false scientism, this forgoing of our responsibility for the formation of public opinion, is apt to decrease the importance of our work for making people more rational.

Unlike his colleagues, however, Myrdal did not enjoy writing for the daily newspapers. “You know my position: I have never written a newspaper article,” he wrote (in an undated letter of 1934) to Heckscher. Still, a few articles did appear by him in time as well as many interviews by the media. Perhaps he
had less of a need to publish in newspapers as he could makes his views known through interviews, government investigations and books once he had reached the position of being a "wonderboy".

COMPARING THE FIVE

The quintet of Wicksell, Cassel, Heckscher, Ohlin, and Myrdal represents Sweden's foremost economists. They demonstrated a wide-ranging repertoire and worked on a great variety of society's problems. For two of them—Heckscher, the economist and historian, and Myrdal, the institutionalist—this breadth of approach was a consequence of their scientific orientation.

Each had a distinct ideological orientation and was at the same time a committed theoretical economist, political ideologue, economic policy expert and debater. Wicksell performed differently in his roles as scholar and political debater—in the one case, unassuming and cautious, in the other, vociferous and bold. He believed that by having adopted a utilitarian position he had found a way of raising scholarship above the clash of value judgments and interests. Cassel admitted that he built on certain liberal values but held that if only economic life were permitted to rest on this foundation, the result would be a growth of prosperity which would benefit the vast majority. On the whole, however, he seems to have regarded himself as the upholder of reason and more or less absolute truth. Of the earlier economists, it was Heckscher who discussed the relationship between scholarly work and politics in greatest detail. He was somewhat pessimistic concerning the scholar's chances of elucidating even for himself where the boundary came between the one and the other, but he believed that the neoclassical theory gave some degree of guidance. Ohlin—thanks to Myrdal perhaps—was conscious of the complications arising from the fact that the scholar is governed by his value judgments but did not always let the world know of his possible cogitations over the question. Myrdal did cogitate more, and landed in a relativist position. Value judgments entwine most things: choice of problem, empirical material, method, and solution.

The three older economists wanted to stand clear of partisan commitments. They certainly had views about party politics—Wicksell both criticized and commended the Liberal and Social Democratic parties, and Cassel willingly drew the guidelines up for a “true” Liberal party and the
curtains down for the Social Democrats (one of his books was entitled *Socialism eller framåtskridande: Socialism or Progress*) but, to cite Cassel’s favorite expression again, they were reluctant to find themselves “popped into a particular pigeonhole,” because that would curtail their freedom of action and thus their credibility. The two younger economists do not seem to have let that matter worry them but entered the political fray with gusto. Ohlin gives the impression of having assumed himself to be a strong enough personality not to have to compromise his ideals. Myrdal, on the other hand, expressed distaste for possibly having, as a member of a collective decision-making body, to restrain his personal views.

All five were broad, not only in that they addressed a wide set of issues, but also by virtue of their endeavors to reach out to the public at large. To speak clearly and simply to the public was in their eyes an imperative duty of the academic economist, and some of them indeed seem to have felt themselves to stand in an intuitive relationship with the “masses”.

The ideological extremes of Cassel and Myrdal display striking similarities. Both held that a good social scientist ought to strive for breadth of approach; both were clear that they did not stand on ground free from value judgments; both sought realism in preference to “metaphysical speculation”; both strove for simplicity and clarity; both believed it their duty to speak to the public at large. The point at issue between them consisted of the differences between their value judgments and of Cassel’s being more of an “absolutist”—he considered his values and views to represent some kind of absolute reason, whereas Myrdal was a “relativist” who thought it possible to argue in reasonable terms from different value judgments.

The five great Swedish economists were undoubtedly supported and inspired by their wives. Anna Bugge Wicksell organized the somewhat chaotic life of Knut and made him focus on scientific production. She actively tried to stop him from writing for the newspapers. She also had her own agenda: she fought for women’s suffrage, was active in the Liberal Party and worked within the League of Nations actually as the first Swedish woman diplomat. Alva Myrdal worked initially closely together with Gunnar but made later in life a career in some ways more sensational than his: she was involved in the development of Sweden’s social welfare system, worked within UNESCO, represented Sweden at the Geneva disarmament conference, was awarded the Nobel Peace Prize and was a Member of

---

23 See Jonung and Persson (1997) for a summary of the “story of Anna and Knut”.

---

ECON JOURNAL WATCH 540
Parliament and of the Cabinet. In the public debate on the population question in the 1930s, two married couples met: on the one side Gunnar and Alva Myrdal, on the other Eli and Ebba Heckscher. (See the appendix for a summary of the personal relationships that emerged among the five economists.)

**SWEDISH ECONOMISTS IN FUTURE PUBLIC DEBATE?**

Wicksell, Cassel, Heckscher, Ohlin, and Myrdal founded Sweden's tradition of media-tuned university economists strongly involved in the current problems of society. They set an example inspiring subsequent generations of economists. Many of today's Swedish economists in academia have followed in their footsteps—often without reflecting much about how and why this pattern originated. More recently, however, there are signs that the opinion held by academic economists, in particular among younger generations, regarding the economist's role in public debate has shifted away from the ideal which this quintet of earlier scholars endeavored to live up to. A number of tendencies are contributing to this shift.

There is today an inclination among academic economists either to withdraw from public debate or else to call debating contributions into question by arguing that “strict” or “rigorous” scientific proof does not exist for policy recommendations. Consequently, the "serious" economist would do best to refrain from expressing himself or herself in public

---

24 Ohlin and Myrdal were followed by a third generation of economists who presented their doctoral theses in the 1930s. Some of the members of this group became prominent civil servants like Dag Hammarskjöld and initially Erik Lundberg. Dag Hammarskjöld left academic life for a career as civil servant. However, as Secretary to the Ministry of Finance 1936-46 and as Chairman of the Board of the Riksbank 1941-48 he was one of Sweden's most influential economists. He was in fact a "politician," although he regarded himself as being elevated above party politics. In public debate, however, he kept a very low profile, leaving no journalistic traces in economics. Erik Lundberg was his opposite; he became a public figure through his position as head of Konjunkturinstitutet from its start in 1937, as publisher of books, articles in the press and in the quarterly journal of *Skandinaviska Banken* in the 1950s, 1960s, 1970s and 1980s. After returning to academia as full professor, he carried on the tradition of Wicksell, Cassel and Heckscher by being highly active and visible in public debate. For a summary of Lundberg's life and contributions, see Henriksson (1994).
debate—in sharp contrast to the relish with which earlier economists hurled themselves into public controversy. One example is a study by Skedinger and Johansson (2004) demonstrating that the most merited economists, full professors conducting research on international trade and capital movements, have been inactive in the debate on globalization. However, their article, which appeared in a Swedish version in 2002, inspired a change and led to greater efforts aimed at the lay public.

In addition, competition for jobs in the increasingly “technicized” and “mathematicized” field of economics requires heavy concentration on academic production, that is, on articles accepted by refereed journals in English. Professional prestige rests today almost solely on the number of articles published and citations received in the professional journals, all printed in English. Thus, writing in Swedish is counterproductive for the young scholar who has not obtained a permanent position.25

The teaching load at the universities makes heavy inroads on the time and commitment of professors as well. Wicksell, Cassel, Heckscher, Ohlin, and Myrdal were operating in a different academic environment, far from today’s mass production of undergraduate and graduate students. In their day, advising and training future doctors of economics was a minute element of a professor’s work as doctoral students were expected largely to look after themselves.26

A growing group of today’s academically active economists, often with an abstruse mathematical specialization, seems to be more isolated from the public debate than earlier generations of economists, while functioning at the same time as trend-setters and role models for young aspiring PhDs and researchers. The researcher who involves herself or himself in burning issues of the day runs the risk of being regarded as less earnest and academically ambitious, besides spending time on publications that will be classified as of no academic merit. The next step may be that polemical articles and debating contributions will be regarded as demerits.27

25 This is seen for example from an article by Lindqvist (2003) which ranked all economists at Swedish universities according to their publications in a number of English-speaking scientific journals. No weights were given to any publications in Swedish. This ranking set off a lively debate within the profession.

26 See Jonung (1991, 1992) for a discussion of the five economists’ contributions to the nurturing of new generations of Swedish economists.

27 If these tendencies persist, who is going to ascend to the throne of public debate from which academic economists have abdicated? Will it be economists associated with banks, brokerage firms and other organizations or will it be representatives of other social sciences than economics? Here political science is at present a major challenger of economics.
If this trend continues, less and less importance will be attached in
the future to the duty and social responsibility of participating in public
debate and imparting some grasp of economic thought to the public at
large. Perhaps we are moving towards a future in which academic
economists have abandoned the social commitment that served as a guiding
principle for the founding fathers of the field in Sweden.

There have been voices raised against this development. As we have
seen, Gunnar Myrdal issued a warning already in the late 1960s. Several of
the contributions in Jonung (1996) express fear for a withdrawal of
economists from public debate.28 Another warning was given by Assar
Lindbeck (2001, 32) who argued that “we do not educate enough ‘two-
legged economists,’ who both master analytical techniques and have a
feeling for real-world problems.” He went on: “This may be a reason for
the receding role of academic economists in the general discussion of
economic and social problems in several European countries.”

Four Swedish PhD students recently conducted a survey with results
Swedish graduate students in economics were “one-legged” when admitted
to the PhD program, and if they risked “losing a leg” during their doctoral
studies. Their conclusions are broadly similar to those reached earlier for
the US by Colander and Klamer (1987). Students entering the program
were “two-legged” but there was a risk they left “one-legged,” at least the
graduate program did not encourage an interest in real world issues.

The study by the four PhD students represents a challenge to the old
tradition of university economists taking part in public debate, but also a
sign of hope. Many of the doctoral students expressed a desire in taking
part in public debate, many were pulled into graduate studies by an interest
in the social sciences, and 17 percent of them had already published a
debate article while being in the doctoral program. This suggests that the
tradition from the great Swedish economists may be surviving.

---

28 This volume was initiated by the complaint raised in the media against the economics
profession for not giving adequate warning of the depression that hit the Swedish economy
in the early 1990s.
APPENDIX: A NOTE ON THE PERSONAL RELATIONS BETWEEN THE FIVE

With such a group of strong personalities moving around in the Swedish academic and political ponds, both friendly and hostile relations developed, and of course there were many incidents of one kind and another. They were all strong-willed individuals and were deeply engaged in debate, public as well as academic. Some spectacular examples demonstrate how messy things can be when personal, scientific, journalistic and political lines and ambitions criss-cross in a small country like Sweden. Hopefully, these examples may also illustrate some similarities and differences between the five.

—Wicksell and Cassel fought like cat and dog. The conflict started when Cassel attacked and Wicksell defended marginal utility theory, accelerated when they competed for the chair in economics at Lund University at the turn of the century, and culminated when Wicksell published a derogatory review of Cassel’s *Theoretische Sozialökonomie*. Wicksell was especially critical of Cassel’s self-assured scholarly attitude.

—Cassel and Heckscher were both, after World War I, classical liberals and fought for many years side-by-side against socialism, economic planning, and the policy views of the Stockholm School and Keynes. But Heckscher was critical of Cassel’s attitude, just as Wicksell was, and the two of them fought a fierce battle over tariffs since Heckscher was a “dogmatic” free-trader while Cassel had a more pragmatic view of the usefulness of tariffs.

—Wicksell and Heckscher had very friendly personal relations. Both Heckscher and Cassel, however, were critical of Wicksell’s tendency to disregard complications of the real world, such as when, after World War I, Wicksell wanted to push the price level back to where it had been prior to the war in order to eliminate the injustices of wartime inflation.

—Cassel was Myrdal’s mentor and, while Cassel was the most right-wing and Myrdal the most left-wing of the five, they retained a close and warm personal relationship over the years. When Myrdal succeeded Cassel as professor of economics at Stockholm University, Cassel embraced him...
and said that he (Myrdal) was the most dangerous man in the country, but at the same time his only worthy successor.

—Heckscher was Ohlin’s mentor and the latter succeeded him as professor of economics at the Stockholm School of Economics. While the political rift between the two was less glaring than that between Cassel and Myrdal the consequences for their personal relationship were much more serious. The relationship collapsed when Heckscher accused Ohlin and Myrdal of being economic planning zealots and (in a separate letter) accused Ohlin of kowtowing to the people in power, that is to the Social Democrats.

—Ohlin and Myrdal were allied at the outset of the battle over unemployment policy, which gradually evolved into a debate over economic planning; their main opponents in this debate were, of course, Cassel and Heckscher. Ohlin and Myrdal formed, with Erik Lindahl, the “older”, neo-Wicksellian nucleus of the Stockholm School—the name that Ohlin gave the school of thought that emerged in Sweden in the early 1930’s. (This concept should not be confused with the business school in Stockholm bearing the same name.) But they parted ways and became political adversaries, Myrdal being a Social Democratic strategist and leading Swedish advocate of the planned economy, Ohlin stepping forward, as leader of the Liberal Party, as the main contender against the Social Democratic hegemony. In the end, consequently, Ohlin had to fight the battle against advancing socialism and economic planning, much as Heckscher had done before him.29

—The list can be made longer. Actually, one reason why the Stockholm School did not develop into an expanding and active body of thought was personal tensions between leading economists in Sweden. See the introduction and summary in Jonung (1991).

29
REFERENCES


Heckscher, E. 1944. Erfarenheter av ekonomi och ekonomisk politik under fyrtio år, *Ekonomien*.


ABOUT THE AUTHORS

**Benny Carlson** (born 1950) is associate professor at the Department of Economic History at Lund University. He started out as a journalist in the 1970s. His main area of interest is the history of economic and social ideas, particularly the exchange of ideas between Germany, the United States, and Sweden during the 20th century. His research has of late revolved around “Sweden and Bretton Woods” and “American welfare capitalism and social engineering.” He is also engaged in research and evaluations concerning immigrants on/off the Swedish labor market. His email address is benny.carlson@ekh.lu.se

**Lars Jonung** has since September 2000 been Research Adviser at DG ECFIN, European Commission, Brussels, dealing with macroeconomic issues. He was previously professor of economics at the Stockholm School of Economics. His research is focused on monetary and fiscal policies and the history of economic thought. His books include *The Long-Run Behavior of the Velocity of Circulation: The International Evidence*, (with M. D. Bordo, 1987); *The Political Economy of Price Controls: The Swedish Experience 1970-1985* (1990); and as editor *The Stockholm School of Economics Revisited* (1991); *Lessons for EMU from the History of Monetary Unions* (with M. D. Bordo); *Bertil Ohlin. A Centennial Celebration. 1899-1999* (with R. Findlay and M. Lundahl as co-editors). He has also published several books in Swedish. His email address is Lars.Jonung@ec.europa.eu.
DO ECONOMISTS REACH A CONCLUSION?

Do Economists Reach a Conclusion On Rail Transit?

TED BALAKER* AND CECILIA JOUNG KIM**

Abstract

FOR MANY POLICY ISSUES THERE IS GENERAL AGREEMENT ON policy goals. That is not the case with rail transit. Depending on who you ask, rail transit should reduce traffic congestion, improve mobility, get motorists out of their cars, clean the air, stimulate the economy, boost property values, concentrate development, fight sprawl, decrease social alienation, help make a city “world class,” and so on. Once we focus solely on economists the list of potential goals shrinks considerably, but we are still left with a bundle of goals.

Summarizing economists’ views on rail transit poses additional challenges. Even simply separating the economists from the non-economists can be quite challenging.

Many researchers from many different disciplines have weighed in on the merits of rail transit. In this case, we will examine only the opinions of economists, that is, those who have at least a Master’s degree in economics or who have taught economics at the college level. In cases of coauthorship,
we include the reference if at least one of the authors is an economist. We exclude the judgments of many of the most prolific researchers in rail transit, notably urban planners and engineers. Indeed, the disagreements over rail often stem from the fact that different researchers from different disciplines think rail should accomplish different goals.

Another challenge comes in attempting to focus on rail transit. Since transit modes tend to operate under the same kinds of incentives, economists may analyze public transit in general (i.e. rail, bus, etc.). We focus on rail transit, but, when appropriate, take a more general view of public transit. Apart from separating rail from other modes, there is also the matter of distinguishing between heavy and light rail. The extensive and well-established systems in places like New York City, Washington, D.C., Chicago, and San Francisco are heavy rail, while newer systems and extensions tend to be light rail.

There are important operational differences between the two types of rail which affect their performance. For example, heavy rail enjoys greater speeds and carrying capacity in large part because it is grade-separated, typically operating underground or on elevated tracks. Light rail is generally at-grade, but not always.

Here we examine U.S. cases and consider rail’s success or failure as determined by the criteria set by each individual economist. And so the big-picture question—Do economists think that rail-transit is good policy?—will be a compilation of judgments given by economists who consider different criteria.

Our general approach is to quote amply and allow readers to decide how much weight to give to the various arguments, studies, and such. However, in some cases readers may benefit from our giving analysis and perspective, so occasionally we do just that. Of course, this introduces the possibility that some of our biases will creep in and so we will make it clear where we stand. The lead author is generally skeptical about rail transit’s success in most areas in the U.S. Whether the issue is highways or transit, rail or bus, he generally favors a transportation policy supported by user fees. Public transportation is often characterized by public subsidies, and here he favors cost effective policies that focus on improving conditions for the poor.

---

1 We do not address high-speed rail and address commuter rail only briefly.
2 In some cases, heavy rail systems do have some light rail components as well.
3 According to the American Public Transportation Association, average operating speeds in 2004 were 20.4 miles per hour for heavy rail and 15.5 mph for light rail. For commuter rail (not the focus of this study) average speeds were 31.5 mph.
TRENDS IN USAGE AND FUNDING

Economists often express concern that we are spending more on public subsidies to transit, but getting less. Baum-Snow and Kahn (2005) examine the 16 major metropolitan areas that established or expanded rail transit infrastructure from 1970 to 2000. They estimate that federal, state, and local governments spent more than $25 billion on construction.

Billions more have been invested to maintain and improve existing rail transit lines. Despite the significant infrastructure improvements associated with these improvements, transit ridership has been declining rapidly. . . . Across all metropolitan areas, the fraction of [those working outside the home] using public transit fell from 12 percent in 1970 to just 6 percent in 2000. These declines have occurred in metropolitan areas with historically high transit use and significant rail infrastructure in 1970 (old-transit cities), metropolitan areas that established significant rail transit infrastructure since 1970 (new-transit cities), and metropolitan areas without rail transit in 2000 (no-transit cities). Though in percentage terms rail transit cities saw less rapid declines in use than cities with no rail transit, transit lost more market share in cities with rail lines. (Baum-Snow and Kahn 2005 1, 4)

The authors see the benefits of rail construction and expansion as flat.

[W]e find little evidence that significant ridership gains due to new rail lines continue to accrue more than a few years after construction is completed. (Baum-Snow and Kahn 2005, 2)

A network-effects argument would expect that newly added lines would spur larger ridership gains because riders would have access to more possible destinations. And yet, “We find evidence of decreasing marginal returns to new rail investments for every city that had rail transit expansions in more than one decade except Portland and perhaps Atlanta” (Baum-Snow and Kahn 2005, 39).
Winston and Maheshri (2006) draw similar findings.

In 1980, two million Americans got to work by transit. Today, in spite of an increase in urban jobs and transit coverage, fewer than one million U.S. workers commute by rail, causing its share of work trips to drop from 5 percent to 1 percent. Although rail transit’s farebox revenues have consistently failed to cover its operating and capital costs since World War II, governmental aid to cover transit deficits has been increasingly available. Since 1980, annual operating subsidies have climbed from $6 billion to more than $15 billion today (APTA Transit Fact Books, figures in 2001 dollars). Capital subsidies have also increased as transit agencies struggle to maintain and provide new facilities, track, and rolling stock. (Winston and Maheshri 2006, 2)

Even in the early twentieth century world events and local policies that governed private transit monopolies were already causing financial strains. Hilton (1985) summarizes conditions in 1918.

The industry had an extremely inflexible pricing structure with a five-cent fare enforced nearly everywhere either by franchise or by municipal regulation. The effects of the draft, expansion of factory employment, and the strength of the union, the Amalgamated Association of Street Railway Employees, caused labor costs to rise by 85 to 90 percent in 1918. Many of the basic materials used by the industry rose in price by about the same amount, and the price of inputs used also for military purposes increased even more (for example, asbestos by 560 percent). The industry’s output declined both absolutely and relative to population. (Hilton 1985, 38-39)

As auto travel grew, transit ridership suffered. In 1964 the federal government stepped in to attempt to rescue transit patronage. Yet, according to Charles Lave, federal subsidies brought new problems.

---

4 Does not include commuter rail ridership.
Consider the urban transit “problem.” In the 1960s the problem was declining transit patronage. Finances received little discussion because the industry was essentially self-supporting: operating costs were so low that passenger revenues covered costs. In the 1990s “problem” has a whole new meaning: financial deficits. Today, most transit revenue comes from governments, not passengers, and the result is continued fiscal crisis—the search for money to continue the subsidies. (Lave 1994, 21)

Subsidies were designed to save transit, but they lead to severe deficits and a large decline in productivity, that is, output per dollar of input.

Indeed, if transit productivity had merely remained constant since 1964, when federal intervention began, total operating costs would be more than 40 percent lower—enough cost reduction to erase most of the current operating deficit, without raising fares. It is uncommon to find such a rapid productivity decline in any industry. In general, productivity increases over time, and a given quantity of input produces more and more output—which is why per capita income rises. Thus the productivity change in the transit industry is notable for both its downward direction and its magnitude. (Lave 1994, 21)

The initial federal subsidy program in 1964 was confined to capital subsidies, but in 1974 the Urban Mass Transit Administration began subsidizing operating costs as well.

[T]he decline in productivity accelerated by 50 percent when capital subsidies began and accelerated by another 48 percent when operating subsidies were added as well. . . . Our attempts to solve the original problem created a new one: a serious decline in the basic productivity of the industry, with an inevitable growth in financial problems as the result. (Lave 1994, 24, 25)
BIASED PROJECTIONS OF RIDERSHIP AND COSTS

Economists have long been critical of faulty forecasting (Gómez-Ibáñez 1985). Haynes Goddard contends that local decision-makers are unfairly critical of rail. He cites a failed attempt to persuade Cincinnati voters to support a sales tax increase that would have funded a light rail system. Goddard argues that the economic analysis overestimated costs and underestimated benefits.

Local officials decided that a genuine benefit-cost analysis of the proposed starter line should be undertaken (and the subsequent alternatives as well). A public solicitation was held, and a competent economic consultant was retained for the study. . . . The study was very conservatively done in that some benefit categories were excluded, such as development benefits around the stations. . . . On the cost side, the most extreme cost conditions were included in order to ensure against the all too common bane of public investment projects, cost underestimation. An explicit uncertainty analysis was conducted (Monte Carlo) and the probabilities of the outcomes were generated with the aid of a regional panel of economists and corporate financial executives who oversaw the entire process. (Goddard 2004a, 4)

However, Goddard appears to represent the minority view. For example, after conducting 209 interviews with public officials in Southern California and elsewhere Jonathan Richmond concludes that the evaluation process is unfairly tilted to favor rail.

The evidence from the Long Beach case shows that decisions did not follow from technical analysis, but that technical analysis was used to support an already existing predisposition to rail transit. Interviews conducted for this project in San Diego, San Jose, Sacramento, and Portland, Oregon also indicated that decision makers had made up their minds prior to the conduct of technical analysis. Technical analysis, then appears to serve a ritual function: it gives an aura of respectability to decisions which have
been reached on other grounds, rather than focusing on the decisions themselves. (Richmond 2005, 147)

Inaccurate cost and ridership forecasts make others wary of rail:

In most cases, actual ridership has fallen far short of projected figures. This is important because, in many instances, local development agencies use ridership projections to convince the Urban Mass Transit Administration of the Department of Transportation that their rail projects are worthy of government assistance. (Zaretsky 1994, 10)

Economists frequently treat biased projections as an unsurprising reality: “It is very common for transit planners to underestimate construction costs and overestimate future ridership” (Baum-Snow and Kahn 2005, 21).

Winston seems to expect inaccuracies to continue:

Indeed, if experience is any guide, rail ridership will be grossly overestimated at the planning stage for new service, while capital and operating costs will be underestimated. (Winston, 1999a)

Even so, Goddard (2004a, 2004b) charges that critics of rail transit are corrupt:

It is said that when you cannot understand why people take the positions on issues they do, it is usually for at least one of three reasons: ideology, money, or sex. In the case of the various light rail transit critics around the nation, mostly associated with libertarian institutes, ideology certainly is involved, and given their general unwillingness to reveal the sources of their funding for their irrational enthusiasm for highway based transport policies, it is an easy conclusion that money quite likely is involved as well … As Upton Sinclair once quipped, it is difficult to get a man to understand something when his income depends on his not understanding it. (Goddard March 2004, 1)
Yet Goddard’s assessment of rail critics as right wing or libertarian is misleading. Many of the institutions that have produced the greatest amount of rail skeptical research—for example, the University of California system, University of Southern California, the Brookings Institution, and Harvard—can hardly be considered libertarian. Many libertarian researchers are skeptical of rail, but rail skepticism is hardly a bedrock belief of libertarian philosophy. Like economists in general, libertarians place considerable weight on cost-effectiveness and this helps shape their views about many issues, including rail transit. If and when rail proves the best, most cost-effective choice, libertarian scholars are likely to support it. Moreover, libertarian scholars tend to support the “user pays” principle which calls for reducing subsidies in all aspects of transportation policy, from transit to highways. Note, for example, their strong support of toll roads and variable pricing, policies that place the costs of driving more directly on the shoulders of motorists themselves. Lindsey (2006) shows that such views are typical of economists who write on highway management.

Speculation about motivations can be applied to any side of any policy issue and such conjecture offers little illumination. Scholarship should be judged, first and foremost, on its merits. But Goddard charges that much scholarship that is critical of rail is not well done. He views rail critics as not just corrupt, but incompetent as well.

In the absence of a clearly articulated and explicit model for predicting transportation choices, the critics can make cheap and superficial but plausible sounding arguments. But in fact, most of their writing reflects either an intellectual incapacity to formulate and reason from explicit analytical models, or in some cases where the training of the authors (some with Ph.Ds) ought to permit this, the writing borders on intellectual dishonesty. Tellingly, these critics don’t publish in the peer reviewed literature, with few exceptions, because most of what they write would be shredded by professional reviewers. Their

---

5 For example, Robert W. Poole, Jr., the founder of the libertarian Reason Foundation is well known as a pioneer of toll roads and variable pricing. He developed the High Occupancy/Toll (HOT) lane concept in 1987 and the term first appeared in a 1993 Reason Foundation study.
writing certainly would be awarded a failing grade in virtually any economics class. (Goddard 2004a, 3)

Goddard’s assessment of rail skeptic’s academic credentials is wildly off the mark. He overlooks the many rail-skeptical economists who frequently publish their findings in peer reviewed journals, write books published by academic presses, and hold positions at major universities. (The references cited in this article provide a sample of this, including articles published in top-tier journals such as the *Journal of the American Planning Association* and the *Journal of Urban Economics.* Indeed, one of the most prolific critics of rail transit, the late John F. Kain, was chairman of the economics department at Harvard. Most would consider such a man eminently qualified to conduct economic analysis.

One of the earlier systematic critiques of rail transit was written by Donald Pickrell of the U.S. Department of Transportation. Pickrell (1990) examines ten rail projects and discovers that only in Washington, D.C. was ridership more than half of what was forecasted. “The consistent over-estimation of future ridership on recent rail transit projects suggests that, with few exceptions, the levels of travel and related benefits they currently provide are far below those originally anticipated by the local decision-makers who selected these projects” (x). Pickrell found a similar pattern for cost estimates. “[C]aptial outlays for Pittsburgh’s South Hills light rail reconstruction project were actually 11 percent below their forecast value, while cost overruns on other projects ranged from 13 percent for Sacramento’s recently completed light rail line to 106 percent for Miami’s downtown ‘Metromover’ project” (xii). Sacramento’s operating expenses were lower than expected, but elsewhere “actual expenses range from 12 percent to more than 200 percent above their projected levels” (xiii). In the *Journal of the American Planning Association*, Pickerell (1992) reviewed the finding in an article titled, “A Desire Named Streetcar: Fantasy and Fact in Rail Transit Planning.”

Richmond (2001) recalls the controversy that followed Pickrell’s report:

The transit industry responded angrily. . . . The [American Public Transit Association] was particularly concerned that ‘The report used projection data made during the very early project planning stages, rather than in the revised and more accurate submitted with the Final Environmental Impact Statement or the Full Funding Agreements.’
Balaker and Kim

Pickrell had made clear, however, that since subsequent (and generally less optimistic) forecasts made after the decision to proceed were by definition irrelevant to the choice of rail, ‘this study focuses upon the accuracy of projections that were available to local decision-makers at the time the choice among alternative transit improvement projects was actually made’ (p.3 emphasis in the original). (Richmond 2001,149)

Richmond defends the Pickrell report, calling it “a model of clarity.” The conclusions in Richmond’s 2001 analysis of rail in twelve U.S. metropolitan areas are very similar to those reached by Pickrell.

Optimistic claims that new urban rail systems would increase transit ridership, reduce congestion and improve the environment while at the same time improving the financial performance of transit systems have proved incorrect in most instances examined here. The evidence shows that the capital funds spent have generated few benefits. While rail’s contribution to increasing transit ridership on the systems under review has been mostly minimal, changes in bus operating practices designed to accommodate rail have generally had a negative effect on the financial productivity of the transit systems concerned. (Richmond 2001, 172)

Kain analyzes forecasting problems in locations such as Los Angeles (1988), Dallas (1990), and Houston (1992). Lave (1991) suggests introducing self-correcting discipline in the forecasting business, in which consultants would have to bond their forecasts. In cases of significant errors, Lave argues that consultants should be forced to give back their fees or build and operate the rail system at the cost they initially projected.

When transit patrons move from buses to service offered by a new rail line, the substitution can add an additional challenge for those who examine ridership figures. Since patrons who switch from bus to rail are not increasing overall transit ridership, modal substitution may at least partially offset rail’s ridership gains.

Analyses contained in this paper also demonstrate that widespread claims that [Atlanta’s] MARTA achieved large
increases in transit ridership by building rail are incorrect and result from the mistaken use of boardings rather than linked trips to measure system ridership. When MARTA implemented its rail system, it modified its bus network to feed the new rail system. This practice increased rail boardings, but more importantly dramatically increased system transfer rates. Many trips that previously had been made as a direct bus trip with one boarding required two or three boardings after the bus feeder-rail network was implemented—a bus-rail, rail-bus, or in some instances a bus-rail-bus trip for the same journey. Use of boardings thus greatly overstated the increases in ridership that actually occurred during the period of rail operation. (Kain 1996, 23)

Reviewing the earlier decades of rail, George Hilton (1976) concluded: “In general only about 8 to 12 percent of the riders on a rapid transit line are former drivers of automobiles. Typically, more than 80 percent are former passengers on bus lines or pre-existing electric railways.” Lave (1998) puts the figure at 70 to 75 percent and Richmond (2001) reports that the percentage of rail riders who were former bus riders ranged from about a quarter to three-quarters. Baum-Snow and Kahn (2005) find that most rail riders are former bus riders, yet while commenting on that research, Voith suggests that widespread mode-switching is not necessarily a shortcoming of rail.

[I]t is not clear that in the absence of the rail investment bus riders would have remained transit users over time. The authors argue that the differences-in-differences approach effectively controls for trends pre- and post-rail investments, and therefore controls for trends in bus ridership. However, the populations in neighborhoods evolve over time, and there is little reason to expect that the people living in neighborhoods ten years ago are the same individuals (and certainly not in the same point in their life cycle) as those in the neighborhood ten years in the future. It seems highly unlikely that the demographic characteristics of ridership on the Washington Metro would be the same as the characteristics of those riding the
bus in the city if the metro had not been built. (Voith 2005, 56-57)

Kain and Liu (1995) argue that rail can compromise ridership gains attained via service increases and fare reductions.

Regrettably, in all three cases [Portland, Los Angeles, and Atlanta], regional transit authorities abandoned their highly successful policies of increasing service levels and reducing real fares for policies that entailed using growing shares of available subsidy dollars to build and subsidize the operations of costly and ineffectual rail systems. Frequently, the introduction of rail services was accompanied by increases in real transit fares. (Kain and Liu 1995, 623)

MOBILITY AND CONGESTION RELIEF POTENTIAL

A key issue that many economists examine is whether rail transit is cost-effective in reducing traffic congestion. One way economists judge this is whether rail can lure commuters out of their cars. On this count, Baum-Snow and Kahn (2005, 51) point to rail systems in Boston and Washington as “standouts” compared to others that have been significantly expanded since 1970. It appears that the authors are most encouraged by Washington.

Washington in particular stands out as a city in which commuters are significantly better off as a result of having the option of using rail transit. While there are measurable welfare benefits of new rail lines in other rail transit cities, they appear to be much smaller. (Baum-Snow and Kahn 2005, 51)

Lewis and Bekka (2000)6 argue that, especially in the absence of road pricing, rail transit can be a good way to deliver congestion relief. They find net benefits, for example, from rail lines in Sacramento and elsewhere. And,

6 At the time this paper was published David Lewis was President and Khalid Bekka, Vice President of HLB Decision Economics, an architectural-engineering-consulting firm.
again, Washington, D.C., is highlighted. The authors calculate the travel time index for Washington, D.C.’s I-270 corridor to be 1.72, meaning that peak-period travel times exceed off-peak by 72 percent. Without the rail line the authors estimate that peak-period travel times would exceed off-peak by 98 percent. This one branch, they argue, saved 4 million hours of delay in 1999. It also saved an estimated $62 million in fuel, time and other driver costs, while costing taxpayers only $25 million.


[I]t seems to us that transit advocates in the Washington metropolitan area are on solid ground when they cite rail transit’s effect on highway congestion. The bus system alone generates little in the way of driver time savings, both because buses and cars compete for scarce roadway space and bus users who switch to rail do not affect drivers. Of the benefits to drivers, about two-thirds are congestion-related; the rest are parking search costs.

The motorist benefits are dwarfed by the total benefits of the transit system, $2.3 billion per year. [T]raveler benefits from the weekday bus system are $975 million per year, or $7.57 per bus trip in 2000 dollars. Weekday rail produces about $833 million in traveler benefits and a per rail trip welfare benefit of $5.16. Taking into account the regional subsidies for the two modes, these per trip net benefits drop to $5.81 and $4.51 for bus and rail respectively. (Nelson et al. 2006, 12-13)

Yet economists are not in agreement about rail in Washington. Winston and Maheshri (2006) estimate that costs exceeded benefits by more than $200 million in 2000 and Downs sees little in the way of congestion relief.

There is no evidence that new fixed-rail public transit systems in the Washington and San Francisco Bay areas have diminished peak-period congestion on any expressways there. (Downs 1992, 29)
Some argue that rail has certain advantages that allow it to operate efficiently and reduce congestion.

Compared with single-occupant automobiles, public transportation, especially rail systems, is a much more efficient way to move people around a metropolitan area. (Zaretsky 1994, 1-2)

As for reducing congestion, one full 40-foot bus (about 70 passengers including standees) is the equivalent of 58 cars with an average of 1.2 passengers per car. This one bus is the equivalent of a line of autos that stretches six city blocks for traffic moving at 25 miles per hour. Comparing autos and heavy rail, where one full heavy rail car can accommodate about 180 people including standees, a train of six rail cars, holding about 1,080 passengers, is the equivalent of 900 automobiles. Thus, one full six-car heavy rail train is the same as a line of moving cars that stretches 95 city blocks for traffic operating at 25 miles per hour.

As these statistics clearly show, public transportation is energy-efficient and capable of reducing congestion. (Zaretsky 1994, 3)

However, others say that such analysis confuses carrying capacity with actual use. Rail transit has the potential to operate efficiently, they argue, but only if rail cars are filled with passengers. Nearly empty rail cars are not likely to achieve desired goals, such as reduced congestion.

During the mid-1990's rail filled roughly 18 percent of its seats with paying customers, while bus filled roughly 14 percent. In contrast, about one-third of auto's carrying capacity is typically filled. These differences in capacity utilization have clear implications for mass transit's cost competitiveness with autos. While transit's average operating costs per seat mile are lower than auto's, this potential cost advantage is never realized in practice.

---

7 The FTA no longer requires transit agencies to report load factor figures.
because empty seats drive its operating costs per passenger mile much higher than auto's. (Winston, 1999a)

Although it apparently does not affect his opinion of the efficiency of rail transit, Zaretsky (1994, 9, 19) acknowledges that use of rail transit often falls well below capacity: “Even in our nation’s most congested areas, usage levels do not reach much beyond 50 percent of workers. . . . [E]ven light rail success stories are subject to the reality that over all usage levels are low.”

Winston and Maheshri (2006) do find that rail systems offer substantial congestion cost savings, but those benefits are overshadowed by other costs. Although Garrett (2004, 9) contends that “there is little evidence that rail transit has reduced traffic congestion,” he writes that evidence suggests “that light rail may have slowed the growth in roadway congestion in some cities.”

Before light rail was introduced in Baltimore, the roadway congestion index increased an average of 2.8 percent a year. After light rail, however, the index increased an average of 1.5 percent a year. Average annual index growth in Sacramento before light rail was 4.5 percent and 2.2 percent after light rail … For St. Louis, the average annual congestion index growth before and after light rail was 0.89 percent and 0.86 percent respectively. The roadway congestion index growth in Dallas remained at an annual average of 2.3 percent before and after light rail was introduced. (Garrett 2004, 9)

Yet the increase in transit ridership in general (let alone light rail ridership in particular) relative to the increase in roadway travel tends to be too small to affect congestion. For example, since the start of light rail operation, transit has captured only 0.7 percent of new travel in the St. Louis urban area and only 0.2 percent in the Baltimore urban area. Further, it is important to consider alternate congestion relief methods, for without such context one cannot assess opportunity costs. (Below we examine other congestion reduction options.)

Balaker and Staley (2006) point out that in some rail areas it is official policy to encourage traffic congestion to bring more riders to transit. They also reference Minneapolis’s Hiawatha light rail line to show how at-grade
rail lines can increase congestion in the corridors where they are implemented. Economists are generally not encouraged by rail transit’s ability to get motorists out of their cars.

This marginal attraction of drivers is so small relative either to the growth of traffic on roads or to the daily variance of vehicle counts on the freeways that it cannot be perceived. Rail transit lines built in Chicago, Cleveland, and Boston with funds of the Urban Mass Transportation Administration had uniformly shown an ability to take off the roads approximately the equivalent of six months’ to a year’s secular growth in vehicle counts on the parallel highway facilities. This cannot typically be observed relative to the daily variance in vehicle counts. (Hilton 1976)

Winston adds cost considerations.

Rail transit does divert some traffic off the roads, but rail systems are so expensive to build and operate and transport such a small share of travelers that they can be justified on cost-benefit grounds in only a few U.S. cities. (Winston 2005)

Giuliano and Small (1995, 203) put it bluntly: “[I]n the U.S investment in rail transit has proven to be a terribly inefficient way to divert trips from automobiles.”

Rail transit’s average operating speeds are typically well below automobile speeds and having to schlep oneself and effects to and fro and to make transfers typically render transit journeys even longer. Glaeser and Kahn (2003) note that, on average, transit commutes take nearly twice as long as car commutes. Garrett (2003-2004, 4) also cites trip durations: “Given the opportunity cost of time, especially during work hours, it is expected that many people choose not to ride rail transit.”

Fielding (1995, 239) cites additional factors to explain why rail has difficulty attracting large numbers of motorists: “The automobile provides a faster, more convenient and comfortable alternative for most travelers.” Giuliano and Small (1995, 220) contend that rail struggles against the automobile even when large investments in it are made: “Attempts to lure
people into other modes, even with grandly expensive rail transit systems, are impotent against the convenience of individualized motor vehicles.”

Lave agrees and adds that the preference for auto travel is not merely an American quirk.

Attempts to lure people out of their cars failed because they ignored the commuter’s desire for fast, safe, reliable, on-demand, door-to-door travel. And this is not just an idiosyncratic preference of Americans. As Europeans have gotten richer, they have also bought and used more cars. (Lave 1998, 10)

If rail is not adept at getting motorists out of their cars it is unlikely to provide congestion relief. “[Peak-hour] congestion did not decline for long in Portland, where the light rail system was doubled in size in the 1990s, or in Dallas, where a new such system opened, or anywhere else that light rail systems or even new subways have been promoted as antidotes to peak-hour road congestion” (Downs 2004).

Staley (2006) explains why congestion often increases, even when transit captures a considerable share of travel and auto trips decrease.

Suppose a suburban community of 7,000 people generates 2,000 automobile trips of equal distance (so we can hold vehicle miles traveled constant). The city’s population doubles to 14,000, but its boundaries stay the same, so density doubles. This would double the number of commuter trips to 4,000. If the city is well served by mass transit—everyone lives within a quarter mile of a bus, van, or rail transit stop—research suggests automobile trips might fall by as much as 30 percent, or 600 trips. The community would still generate 1,400 automobile trips even though vehicle miles traveled per person falls. If road capacity doesn’t increase, or fails to keep pace with the increase in travel demand, congestion increases. Compact and higher density development becomes congestion-inducing development. (Staley 2006, 77)

According to Small (2005, 10), congestion relief is not among the benefits produced by transit. “While many recently built transit systems
have achieved some desirable effects, none have seriously lessened traffic congestion.”

Baum-Snow and Kahn (2005, 46) find “scant evidence that rail lines have reduced . . . congestion externalities.” Richmond (2001, 160) offers a similar assessment: “In no case has new rail service been shown to have a noticeable impact upon highway congestion.”

Richardson and Gordon (2000, 11) suggest that larger demographic trends have helped tame congestion: “The moral is that it is continued metropolitan decentralization not transit investments (especially rail transit) that keeps congestion under control.”

DECONTROLLING ROAD-BASED TRANSIT ALTERNATIVES

Often, improving mobility and reducing congestion are among the most important justifications given for building rail transit. With these goals in mind, many economists have insisted that mobility and congestion problems arise because the market is not allowed to function. Economists are very supportive of road pricing (for a review see Lindsey 2006) and pricing may help spur transit ridership (Fielding 1995). Additionally, many economists argue that, rather than build rail transit, government needs to decontrol entry into road-based transit, including buses, minibuses, vans, shuttles, taxis, share-ride taxis, and on-the-spot carpool systems (Roth and Shephard 1984; Klein, Moore, and Reja 1997, 94-125; Savage 1999; Winston and Shirley 1998, 89-106; Kain 1999, 388-393; Utt 2003, 9-10).8 This literature is about road-based transit, but the potentialities are often discussed favorably in comparison to rail. In most states, government restrictions are very heavy, eliminating virtually all possible services except for the government planned bus services, airport shuttles, and the highly controlled taxi industry.

8 For a review of economists’ views on taxi deregulation see Moore and Balaker 2006.
OTHER JUSTIFICATIONS FOR RAIL TRANSIT

Rail transit is often justified on grounds not directly related to transportation, such as its ability to achieve various environmental goals. Zaretsky (1994) argues that rail transit’s large carrying capacity allows it to save energy and clean the air. Yet, as noted above, others counter that one must separate capacity from actual use. Just as they are unlikely to reduce congestion, nearly empty rail cars are unlikely to reduce energy consumption or improve air quality.

[S]ignificant pollution reduction from light-rail transit may not be realized for several reasons. First, there is little evidence that rail transit has reduced the number of vehicles on the roadways. . . . Second, large-scale improvement on pollution, assuming no growth in traffic congestion, can only be had if light-rail passengers substitute rail transit for auto transit. (Garrett 2003-2004, 4)

Shapiro, Hassett, and Arnold move beyond theoretical considerations of capacity and, although they address public transportation in general, they find substantial environmental gains.

The facts are clear and indisputable. For every passenger mile traveled by Americans, public transportation consumes about one-half the fuel and energy of private automobiles, SUVs and light trucks. For every passenger mile traveled by Americans, public transportation produces only five percent as much carbon monoxide, less than 10 percent as much volatile organic compounds, and little more than half as much carbon dioxide and nitrogen oxides. Greater use of public transportation offers the most effective strategy available for achieving significant energy savings and environmental gains without imposing new taxes, government mandates or regulations. At current levels of use, every year public transportation saves close to one billion gallons of gasoline and reduces harmful emissions by millions of tons. Increasing Americans’ use of public transit would produce even greater benefits for
our nation’s economy, security, and environment.  
(Shapiro, Hassett, and Arnold 2002, 33)

Still, it seems that most economists are skeptical about rail transit’s environmental benefits. Richmond asserts that new rail service has not had a “noticeable impact” on air quality (2001, 160) and Baum-Snow and Kahn (2005) also find little evidence that rail transit has cleaned the air. Glaeser and Kahn (2003) point out that air pollution has decreased significantly in recent decades. They credit the “greening of the automobile” (38), in which improved technology reduces vehicle emissions even as vehicle miles traveled increase.

Lave (1978) finds that rail systems are net users of energy, partly because of the large amount of energy invested during construction. Winston and Maheshri (2006) also reference the high use of energy during construction.

For instance, Tri-County Metropolitan Transit Agency claims that under the best case scenario, the proposed north light-rail line in Portland, Oregon would save the equivalent of 7,875 gallons of gasoline per day. But the agency also calculates the energy cost of building the line to be 32 million gallons of gasoline. Thus, even using the most optimistic estimates and assuming no depreciation of the capital stock, it would take a minimum of 15 years to even begin to achieve energy savings—and concomitant reductions in emissions—from this rail line. (Winston and Maheshri 2006, 16)

Lave (1978) calculates that San Francisco’s new BART system would generate energy savings only after 535 years of operation and only under a scenario that has not materialized: that auto fuel efficiency would remain at 14 miles per gallon and that the average commuter trip length on BART would not be greater than the car trips it replaced.

Winston and Maheshri (2006) further reference rail’s low load factors, high consumption of electricity (whose generation creates pollution), and the fact that a large share of patrons keep older, high emission cars to drive to suburban rail stations instead of driving newer, cleaner cars to work. They

---

9 The 2006 Transportation Energy Data Book list energy intensities (Btu per passenger mile) by mode: bus transit 4,160; car 3,549; rail transit (light & heavy) 3,228; vanpool 1,401.
assert that “a greater share of rail ridership has, at best, an ambiguous effect on the environment” (Winston and Maheshri 2006, 16).

Lave (1998, 9) argues that “in terms of energy policy, the money spent trying to boost rail and bus ridership was not just wasted, it was actually counterproductive.” Instead he advocates “The Law of Large Proportions,” which holds that, when trying to reduce energy consumption, it is best to focus on the large components of energy use.

[O]nly 3 percent of passenger trips are made on public transportation; cars carry most of the people and use most of the energy. It makes little sense to spend much effort on increasing the modal share of transit, but even a tiny improvement in the energy efficiency of cars will be consequential. An increase of only 0.2 mpg, for the average mid-1970’s car would have more impact than doubling transit patronage. The Law of Large Proportions says the effects of transit policies would have been insignificant even if they had been able to lure people out of their cars. (Lave 1998, 11-12)

Often researchers and public officials justify rail projects, at least in part, on grounds that they spur economic development. In their analysis of 16 cities over a 30 year time span, Baum-Snow and Kahn (2005) discover that mean real household income in new rail transit-accessible areas was below that of other areas. Moreover, except for Atlanta and Miami, the income gap widened after new rail lines opened. The authors suggest that this supports previous conjectures that public transit is a “poverty magnet” (Baum-Snow and Kahn 2005, 25).

Garrett (2004) considers light rail’s effect on St. Louis property values and examines whether the benefits of greater accessibility outweigh the negative nuisance effect (e.g. noise, increased traffic congestion, safety concerns). He finds that within 1,460 feet of a MetroLink station, home values increase the closer they are to a station. But beyond 1,460 feet home prices rise with distance to station:

However, the increase from being farther from the station is less than the increase from being closer to the station. This suggests that for the entire sample, a house will show a premium the closer it is to a MetroLink station …This positive accessibility effect outweighs the negative nuisance
effect; so, the net effect of MetroLink on property values is generally positive. (Garrett 2004, 21, 23)

Weinstein and Clower (2002) study areas located outside the Dallas central business district and within one-quarter mile radius of a light rail station. In general, the Dallas area experienced a strong property market during the study period.

Nonetheless, proximity to a current or future DART LRT station appears to have had an additional positive impact on median valuations for most classes of property (emphasis in original). (Weinstein and Clower 2002, 6)

Weinstein and Clower report that residential properties near DART stations increased 32.1 percent versus 19.5 percent for the control group areas.

Voith (1991) finds a small positive relationship between median home values and proximity to commuter rail stations. Bowes and Ihlanfeldt (2001) reveal that an intermediate distance from rail stations (one to three miles) is where property values tend to be higher, compared to properties within a quarter-mile of a station and those more than three miles away. Yet others suspect that the positive effects of light rail “will diminish rapidly beyond the quarter-mile radius” (Weinstein and Clower 2002, 2).

In their meta-analysis, Cockerill and Stanley (2002) report that light rail enhanced residential property values 2 to 18 percent in Portland, Sacramento, San Diego, and Santa Clara. But again, it was determined that properties located too close to a station (less than a quarter-mile) may suffer nuisance effects and lower property values.

As expected, the capitalization premiums appear to be higher with the establishment of heavy and commuter rail systems. Turning to the light-rail systems, Portland and Dallas have smallest scale operations as regards the distance and stations included. But the land market changes in Portland have been more extensive than those in cities with larger systems (such as Miami). A strong Smart Growth planning initiative in Portland and attempts to install a culture of reducing congestion have enhanced ridership and the premium residential owners place on being located near public transit. Authors of studies in
both Miami and Dallas, where the effects of residential properties were actually negative, suggest that the “car culture” of these cities has limited public enthusiasm for light rail and thus dampened property value impacts. The studies of various systems in California indicate smaller, and highly varied, land market effects of public transit. (Cockerill and Stanley 2002, 10)

Garrett (2003-2004, 3) writes, “Research generally finds that rail transit has a positive impact on residential property values, although the impact is relatively small.”

Cockerill and Stanley (2002) note that there have been fewer studies on rail’s effects on commercial property, but they report premiums of 4 to 30 percent in Santa Clara, Dallas, Atlanta, San Francisco, and Washington, D.C.

In cities where both the residential and commercial markets have been analyzed (such as Washington, D.C. and Atlanta), higher premiums occur in the later. Yet the California studies of both light and heavy rail seem to represent some exceptions. In San Diego the commercial benefits appear to be concentrated only in the central business district rather than any suburban office zones. And in San Francisco and Los Angeles few capitalizations have appeared for commercial properties located nearer transit stations. (Cockerill and Stanley 2002, 10)

Lewis and Bekka (2000) examine 2,830 commercial properties located in Washington D.C. and find that, all else equal, the shorter the distance between a commercial property and a metro station, the higher the value of the property. In their study of the Dallas area, Weinstein and Clower report that the value of office properties within a quarter-mile of light rail stations increased 24.7 percent versus 11.5 percent in the control group areas. The positive effect did not seem to extend to industrial properties and “retail properties show no meaningful difference in the change in median values between outlets located near the rail stations and those in the control group” (Weinstein and Clower 2002, 5). The authors note that their results “suggest that DART rail is an amenity-enhancing service most keenly affecting the market values of properties where people live and where there
are comparatively high concentrations of non-industrial jobs—i.e. office buildings.”

In many communities public officials have partnered with developers and business leaders to pursue “Transit-Oriented Development” (TOD). The idea is to encourage economic development, alter travel patterns, and reduce congestion, often by building high-density, mixed-used developments that encourage walking and transit use and discourage driving. Boarnet and Crane (2001) conducted a literature review on TOD.

Surprisingly, there is little credible knowledge about how urban form influences travel patterns. . . . Given the enormous support for using land use and urban design to address traffic problems, it was somewhat surprising...to find the empirical support for these transportation benefits to be inconclusive and their behavioral foundations obscure. Prior evidence on the link between design and travel is difficult to interpret and tells us relatively little about the behavioral nature of the problem and thus provides a weak foundation for policy advice. (Boarnet and Crane 2001, 172)

Guiliano and Small (1995) cite changing demographics to explain why hoped-for land-use effects may not occur.

There is also reason to doubt that the hoped for land-use changes will take place as result of investment in rail transit investment. High-density compact cities developed before the automobile. In North America they were the result of industrialization and the earlier state of transportation technology. The technology of production required agglomeration. Access to the intercity rail network was critical, and streetcar systems shaped residential patterns for workers. None of these conditions exist today, and the continued shift to service and information-based industry suggests that economic activity will be even less ‘place dependent’ in the future. (Guiliano and Small 1995, 204)

Measuring rail transit’s impact on economic growth can be especially difficult because researchers must control for many factors, from market forces to public policies on issues as varied as crime rates, regulatory
climate, and the quality of public schools. Of course the benefits must also be weighed against the cost of public subsidies: “The increase in property values and economic development are subsidized benefits and may not be greater than the subsidy costs” (Garrett 2003-2004, 5).

Garrett suggests that policymakers first consider why economic development is not taking place in a certain area.

The general consensus from the academic literature and the findings presented in this report is that light rail is not a catalyst for economic development, but rather light rail can help guide economic development. Rather than relying solely on light rail to create economic development, city planners and officials should first address a key question: Why is economic development not occurring in a given area in the first place? Possible reasons include relatively high cost to business start-ups, unattractive locations (crime, poor infrastructure) and unnecessary zoning and regulations. Unless these barriers are lowered or removed, the long-run economic development objectives, with or without light rail, will not be fully met. (Garrett 2004, 25)

If rail improves mobility for the transit-dependent, it may be justified on social equity grounds. Baum-Snow and Kahn (2005) confirm that most rail riders are former bus riders, and, as noted previously, some economists regard this as a shortcoming of rail. Yet here the authors view it as the chief benefit of rail transit.

In contrast to the pollution and congestion reductions touted by many rail transit proponents, we argue that the primary social benefit associated with new rail lines is that they may significantly reduce trip times. Given that the majority of rail transit riders are former bus users, mode switching to rail has the potential to represent large aggregate time savings. (Baum-Snow and Kahn 2005, 51)

Such time savings could be achieved when a transit user switches from bus travel, which is typically just as vulnerable to traffic congestion as private automobiles, to grade-separated rail travel, which avoids such congestion. (Since the critical issue seems to be not bus versus rail, but grade-separation, one wonders whether such time savings could be
achieved with exclusive busways.) If such mode-switching benefits poorer bus riders, then rail can be seen as progressive. However, Baum-Snow and Kahn are apparently undecided on the matter.

Our empirical work suggests that there are distributional consequences from expanding rail transit infrastructure. Suburban workers who commute by car are likely to gain little from improved transit, while bus commuters who work in the CBD enjoy large time savings in many cities. Since bus riders tend to be poorer people, this suggests that rail transit expansions are progressive. This is a contentious point that merits future research.

Transportation scholars have argued that an unintended consequence of rail transit expansion is bus coverage deterioration due to budget reallocations to pay for the new transit lines. If this is true and if the poor are more likely to take the bus than rail transit, then transit expansion could be a regressive public policy. (Baum-Snow and Kahn 2005, 50)

Winston and Maheshri are clearly skeptical of rail’s equity benefits.

Supporters of these systems claim that they are attractive on distributional grounds because they contribute to the mobility of low-income residents. But the median income of rail transit users exceeds the median income in the general population. In addition, rail transit systems have difficulty keeping up with and responding to changes in job growth; thus, they are unable to proved the poorest residents access to employment opportunities in outlying suburbs. (Winston and Maheshri 2006, 17)

A fairly common argument is that transit agencies build rail transit systems in hopes of luring comparatively well-off motorists out of their cars and, in the process, neglect bus service which serves a poorer demographic. Gordon and Richardson (2001) criticize inefficient targeting in which “many highly subsidized rail systems, such as that of Washington, D.C., serve large numbers of the middle and upper-middle class” (14).

In their favorable assessment of the system, Nelson et al. (2006, 21) agree that: “The benefits of the D.C. transit system accrue disproportionately to
wealthy travelers, both in terms of economic welfare measures and raw minutes saved while traveling.”

Winston and Shirley remark on additional dimensions of the matter.

Rail systems are clearly responsive to the interests of upper middle-income riders because fares fall and route coverage and frequency increase as their share of ridership increases. In contrast, fares increase and service frequency deteriorate as the share of (lower) middle-income riders increases, and route coverage contracts as the share of lower-income riders increases. (Winston and Shirley 1998, 81)

Castelazo and Garrett (2004) offer a different way to serve the transit dependent.

While providing public transit to the poor does produce tangible economic benefits, the following example suggests that light rail is not an efficient means of providing transportation to the poor. Specifically . . . the money spent on MetroLink in St. Louis can be used to much better effect.

Based solely on dollar cost, the annual light-rail subsidies could instead be used to buy an environmentally friendly hybrid Toyota Prius every five years for each poor rider and even to pay annual maintenance costs of $6,000. Increases in pollution would be minimal with the hybrid vehicle, and 7,700 new vehicles on the roadway would result in only a 0.5 percent increase in traffic congestion. And there would still be funds left over—about $49 million per year. These funds could be given to all other MetroLink riders (amounting to roughly $1,045 per person per year) and be used for cab fare, bus fare, etc. Castelazo and Garrett 2004, 12-13)

Like many economists, Castelazo and Garrett regard bus service as another preferable alternative to rail.

Instead of building light-rail systems to provide transportation for the poor, communities could expand
bus service, offer more express bus routes or expand on-demand services; these would still realize the benefits of providing public transportation to the poor. Although these other forms of public transportation are also cost-inefficient compared to the automobile, fewer inefficient public transportation systems would be less costly to society. (Castelazo and Garrett 2004, 13)

The preference for bus is not just on equity grounds. In what was probably the first comprehensive estimate of the comparative costs of urban transit modes, Meyer, Kain, and Wohl (1965) found bus transit to be generally more cost effective than heavy rail. Guiliano and Small (1995, 204) argue that “rail transit is far more costly to build and operate than bus transit.” Pardue (1976) touts the cost-effectiveness of bus transit and Downs (1992, 45) argues in favor of expanding bus service: “Needless to say, expanding existing public transit systems—especially bus systems—is much less costly than building new fixed-rail systems.”

Kain holds that bus service generally offers lower costs and higher performance.

With few exceptions, academic studies of the cost-effectiveness of alternative modes have found that some form of express bus system, operating on either an exclusive right-of-way or a shared facility, would have lower costs and higher performance than either light or heavy rail systems in nearly all, if not all U.S. cities. (Kain 1999, 384)

Vincent and Roth (2005) comment specifically on a rail proposal for the Northern Virginia region.

Our conclusion suggests that building rail as planned would be a significant misallocation of resources. It fails to maximize the number of new transit trips that can be generated, because too much is being spent to attract each new transit rider. It also takes resources away from other potential transit projects that could better serve the region. For example, for the same budget as rail, our analysis suggests that a high quality bus transit system could be built in the Dulles corridor and in several other corridors in
Northern Virginia. This would attract many more people to transit, serve many more communities, and do more to relieve Northern Virginia’s notorious traffic congestion.

Moreover, even though our options are more cost-effective, they can provide better service than rail, including more seats for passengers, more frequent service, and more convenient service that skips stops and reduces travel time (emphasis in original). (Vincent and Roth 2005, 23)

Certain economists challenge the popular notion that rail is an inherently more appealing mode than bus. McFadden (1974) finds that, if the cost and time of the trips were the same, rail patrons have no preference between bus and rail. Richmond (2005) maintains that fares, schedule reliability, and travel time are the factors most important to travelers. He explains that decision makers frequently overlook the time passengers spend walking to and waiting at transit stops, and that travelers actually place extra weight on time spent engaged in such activities. “With a rail system there will be more transferring, waiting, and walking on average than with a bus system. The same bus can provide both local and line-haul service and—lacking attachment to a fixed track—can serve a wider range of destinations than can be reached by train” (Richmond 2005, 48).

Morgner (1976) and others tout other forms of rubber-tire transit.

Rail rapid transit is probably the worst step Los Angeles could take to improve transportation. There do exist a variety of other, smaller steps that can easily be accomplished with strong and positive results. The most promising alternatives appear to be "para-transit" operations of several types that promise substantially larger benefits at a fraction of the costs of the proposed rail system. (Gordon and Eckert 1976)

Still, not all economists are enthusiastic about bus transit: “bus service actually increases congestion to motorists, especially when it operates on exclusive bus lanes that would otherwise be available to all vehicles” (Winston 2005).
BROAD JUDGMENTS

Determining whether building rail transit is a wise investment can be complicated when success is assessed in comparative terms. An economist may view one system as better than most, but that does not necessarily mean that economists favor rail in general. Consider Kain’s highly critical assessment of Atlanta’s rail system.

In spite of the serious questions the analyses presented in this paper raise about MARTA’s decision to build a costly rail system, it should be understood that Atlanta’s system is almost certainly one of the most successful rail systems implemented in the United States since the end of World War II. Most of the other light and heavy rail systems built in the United States since the end of World War II would fare much worse in analyses of the kind presented in this paper. (Kain 1996, 23)

Economists find that rail transit delivers benefits in a variety of metropolitan areas, but most regard larger social trends as at odds with rail. As noted above, economists tend to regard other options as preferable, such as decontrolling transportation markets, introducing pricing, or improving bus service. We have also examined how they are also critical of forecasts that frequently understate costs and overstate ridership. Economists seem to be skeptical that rail’s benefits outweigh costs. For example, Richmond reports that optimistic claims “have proved incorrect in most instances” (Richmond 2001, 172); Baum-Snow and Kahn provide “pessimistic evidence” about the success of new rail lines (Baum-Snow and Kahn 2005, 49); Garrett (2003-2004, 5) and others make similar assessments: “Given the size of costs relative to the benefits, the creation of light-rail transit systems or the expansion of existing systems in American cities may be difficult to justify.”

Utt (2005) sees few places where rail transit can succeed. “The truth is that for almost all cities and communities, the economics of rail-based transit just don’t make sense.”

In their analysis of 25 light and heavy rail systems in operation between 1993 and 2000, Winston and Maheshri (2006) estimated the contribution of urban rail operation to social welfare based on the demand for and cost of its service.
We find that with the single exception of BART in the San Francisco Bay area, every U.S. transit system actually reduces social welfare. Worse, we cannot identify an optimal pricing policy or physical restructuring of the rail network that would enhance any system’s social desirability without effectively eliminating its service (emphasis in original). (Winston and Maheshri, 2006, 2)

Welfare performance of the largest systems deteriorated over time, and during the last year of their sample researchers found negative net benefits for every system. Rail systems in New York City and Chicago may be able to become socially desirable, but only if they were privatized, which Winston and Maheshri regard as highly unlikely. They add: “Because no policy option exists that would enhance the social desirability of most urban rail systems, policymakers only can be advised to limit the social costs of rail systems by curtailing their expansion” (Winston and Maheshri 2006, 20).

Although Richard Voith is somewhat supportive of rail transit, he acknowledges that most economists are skeptical of its benefits. “The dominant view of economists has been that rail investments generally have been ineffective and expensive, and the benefits do not justify the costs” (Voith 2005, 52).

**RAIL TRANSIT IN A CHANGING WORLD**

Economists seem to agree on the demographic features that make rail more likely to succeed—high population density, high concentration of employment in the central business district, and strong commute flows to and from the CBD. If you need to get from point A to point B and you have a choice between train and bus, you probably prefer the train, but if you need to get to a point three miles from B, namely, point C, and only a bus will take you from A to C, then that beats taking the train to B. Economists point out that, compared to rail infrastructure, buses and other road-based modes are far more flexible in route differentiation and alteration over time (Kain 1988; Giuliano and Small 1995, 204, 211-215; Richardson and Gordon 2000; Zaretsky 1994, 3; Vincent and Roth 2005).

The CBD-type features that favor rail appear in a diminishing share of the American landscape and today less than 10 percent of our nation’s
metropolitan employment is located in traditional central business districts (Winston and Maheshri 2006). Economists often cite related demographic trends like decentralization, the suburbanization of employment and housing, and increasing wealth to explain why it is difficult for rail to succeed in modern society. “Suburbanization of employment and residences has made it less likely for transit to be a feasible commuting alternative” (Baum-Snow and Kahn 2005, 31).

Gordon expands on the same theme:

[Transit] worked best in a world with high concentrations of origins and destinations (where employment and population densities are high) and with large numbers of people too poor to own and operate an automobile. Both conditions have been declining in the United States and in many other developed and developing nations, explaining the decline of conventional transit both here and abroad.

Population and employment densities are falling in most places, and most people and jobs are choosing not to locate near transit stations because collective transportation is inconvenient and expensive in terms of what really counts, people’s precious time.

In low-density settings where origins and destinations are dispersed, transit that best serves high-volume corridors competes poorly. High-capacity rail systems are, thus, inevitably underutilized, ever more expensive and ever more difficult to provide more of. (Gordon 1999)

Kain challenges perceptions that CBDs have been growing in influence.

Central business district employment levels are especially critical to transit ridership, particularly on the new rail transit systems that were designed specifically to serve large projected increases in central city employment. Although data are fragmentary and difficult to interpret, they, nonetheless, indicate that few central business districts experienced significant employment growth after World War II. This conclusion will surprise many
individuals who have equated the construction of new office space within central business districts of large cities with employment growth. Overly optimistic forecasts of central city employment are important contributors to the excessively optimistic forecasts of transit ridership of proposed rail systems. (Kain 1999, 366)

Goddard stands out as one who sees demographic trends, especially lower densities, as actually supportive of rail transit.

Upon a little reflection, we can see that the density argument is 19th century in nature and therefore antiquated. When most people in the urban areas didn’t own a horse and carriage and therefore walked to their destinations or to the street car stop, the density argument was valid and important. But in today’s world, in the U.S. at least, most people own vehicles, so population density has lost its force. The rising interest in light rail today is caused by highway congestion, and since park and ride facilities are typically part of light rail transit investments, the population density argument is no longer important—what is important is vehicle density on the roadways, which is high and growing higher. This is another way of referring to congestion, and reduction in congestion is the major source of light rail transit benefits. . . . [O]ne can make the interesting hypothesis that light rail transit benefits will always be higher the lower the suburban population density, as that implies more sprawl, higher vehicle ownership and use and consequent higher road congestion. (Goddard March 2004, 3, 5)

Still most economists agree that today’s housing and travel patterns do not fit well with rail transit. Zaretsky contends that it is increasingly difficult to construct effective public transportation systems.

Designing an effective public transportation system is more difficult than it used to be, as urban and suburban development has altered commuting patterns. Today, most commuting patterns are not from the suburbs or urban residential areas to a central business district, but are
instead between points in the suburbs, as businesses have followed their workforces from downtowns to the suburbs. Although this change does not affect bus service much, as routes can be altered to match commuting patterns, rail service can be rendered obsolete because fixed routes cannot be altered quickly or cheaply … If the corridors chosen for rail construction do not match either the commuting patterns of workers or the areas amenable to future industry expansion, cities will end up with little more than expensive tourist movers paid for with government dollars. (Zaretsky 1994, 3)

Some trends may temporarily boost transit, but they are still not strong enough to reverse transit’s declining influence.

Although a few trends, such as increased immigration, have led to temporary increases in ridership in some communities, the complex industrial, demographic, and land use changes affecting U.S. society continue to erode ridership, even among the most dedicated groups of transit users. Soon, the losses will outweigh the gains. (Rosenbloom and Fielding 1998, 2)

Gómez-Ibáñez considers the financial strain that changing demographics puts on transit systems.

The basic problem is how to maintain or increase public transport ridership without widening the gap between transit costs and passenger revenues; this gap, which we call the transit deficit, is usually financed by the taxpayer. In the last several decades most metropolitan areas have been promoting public transportation as a way to reduce traffic congestion, control air pollution by automobiles, and preserve the mobility of citizens who do not have ready access to automobiles. At the same time, however, the underlying economics of mass transit have been deteriorating, as real incomes rise and jobs and residences move to the suburbs. Rising real incomes make labor-intensive services such as transit more costly, and suburbanization produces a dispersed pattern of travel that
is difficult for transit to serve. The result has been rapidly increasing transit deficits with disappointingly small ridership gains (Altshuler 1979; Meyer and Gómez-Ibáñez 1981; Wachs 1989). The debate over transit seldom faces up to these realities. The public has little grasp of the difficulties the local transit agency faces in trying to hold on to ridership despite rising incomes and suburbanization. Transit's growing deficits are often blamed instead on management or labor inefficiencies, accusations that are only partly true. Thus the public is misled into believing it is possible to have it all: better service and ridership without continually rising deficits. (Gómez-Ibáñez 1996, 30)

It is often assumed that the shift to automobile travel occurred during the post-World War II years and in conjunction with the development of the interstate highway system, which was initiated in 1956. Yet the shift had begun much earlier. Hilton (1985) notes that patronage of demand-responsive private jitneys peaked around 1915, but the trend was soon squelched.

Municipal governments proved to be unanimously hostile to this development. They were directly dependent on the street railways for tax revenues and they achieved some implicit benefits from the monopolistic status of the street railways. . . . More basically, the electric streetcar had produced an urban pattern that most municipal governments liked. (Hilton 1985, 37)

“By 1915 the jitneys had given every impression of being a viable industry,” but they were “eradicated in most cities” before America entered World War I in 1917 (Hilton 1985, 37, 38). Freewheeling private transit struggled under unfriendly policies, yet in the years leading up to World War II private automobile travel was already widespread.

The ubiquity of the automobile was producing changes in the urban pattern to which public transit, as it was organized, could not respond effectively. The old pattern of strip shopping streets leading to a centralized business district of office employment, an amusement-restaurant
complex, and a specialized shopping area was slowly being replaced with one of free placing economic activity, based on expectations that most people would arrive by automobile and goods would arrive and depart by truck. (Hilton 1985, 43, 44)

As wealth increased, more Americans chose to drive. “For the majority of society . . . the tendency to turn away from public transit with increases in income was dominant” (Hilton 1985, 44).

During World War II transit ridership was buoyed by policies that restricted automobile production and rationed gasoline. But after the war, increased wealth meant that private automobile travel would continue to overshadow transit. In their analysis of the 1970 to 2000 period, Baum-Snow and Kahn (2005, 31) note that cost of owning an automobile fell. Moreover, “wages have been growing, increasing people’s value of commuting time, thereby making them less likely to commute by bus or rail lines that are slower than driving.”

Hilton expands on that theme:

The automobile has a strongly positive income elasticity of demand, which is to say that as people’s incomes increase, they want more of the services of automobiles. Indeed, they want additional services of automobiles more than proportionally to their increases in income. In contrast, the income elasticity of demand for transit is positive only in relatively low income brackets, below $4,000 per year. In higher income brackets, income elasticity is negative; as people’s incomes increase they turn away from transit to the automobile. The cross elasticity of demand between the two, which is to say the responsiveness of consumers of one service to a change in the price of the other, is so low as closely as to approach zero. Typically it cannot be picked up by ordinary econometric techniques. (Hilton 1976)
WHY DOES RAIL TRANSIT REMAIN POLITICALLY POPULAR?

Despite criticism from economists, rail transit remains politically popular. Since the early 1980s roughly two dozen American metro areas have built or expanded rail transit systems. Many more angle for federal dollars to break ground on more lines. Some economists, such as Hilton (1973) underestimated the power of politics. “By 1980, at the latest, the present rapid transit movement will be looked upon as unsuccessful, misguided, and purely wasteful.”

Supporters often point to rail’s political popularity as evidence that rail transit is good public policy. Why is there such a disconnect between the economic evidence and political reality? Many economists have confronted this question.

Despite the pessimistic evidence we have presented about the likely success of new rail lines, they are being built at historically high rates. Why is this? An important reason is that most of the cost is covered by the federal government. The nature of federal funding is such that they fund capital intensive transit projects like fixed rail over other types of projects that might actually draw more new riders. (Baum-Snow and Kahn, 49)

Among other factors, Winston and Maheshri (2006) cite increased federal earmarks and federal legislation that sets aside for transit 20 percent of revenues from gasoline tax increases. Yet Voith (2005) notes that locals may be willing to shoulder much of the cost themselves.

[Many regions have adopted specific local taxes to support the development of rail transit systems. Dallas, Salt Lake City, and St. Louis are examples of regions that have adopted local taxes to fund significant shares of rail transit investment. Thus in some metropolitan areas the investment is not simply local areas opting for rail because it has a low tax price. (54)]
Castelazo and Garrett (2004) suggest that rail is politically popular because the benefits are “highly concentrated, while the costs are widely dispersed.”

The direct benefits of a light-rail project can be quite large for a relatively small group of people, such as elected officials, environmental groups, labor organizations, engineering and architectural firms, developers and regional businesses, which often campaign vigorously for the passage of light-rail funding. These groups would benefit from light rail, not from the subsidization of cars and money to all potential riders of light rail. (Castelazo and Garrett 2004, 13)

Winston and Maheshri continue the theme of concentrated benefits and dispersed costs.

Why do existing systems continue to expand and new systems get built despite rail’s negative contribution to social welfare? Rail transit enjoys strong support from urban planners who wish to discourage auto use, from suppliers of transit capital and labor, who receive economic rents, from civic boosters, who perceive that a rail system adds prestige to their city, and from city officials, who support investments in a transit system that serves the downtown core because it may help the downtown remain vibrant or keep it from decaying. Until recently, the public has rarely rebelled against the actual costs of new systems or system extensions. In fact, opinion polls suggest that a majority of residents in a city tend to support rail transit regardless of whether they actually use it on a regular basis. We speculate that the public may support rail transit because it overstates rail’s ability to mitigate automobile externalities and because it is “rationally ignorant”—that is, the costs of transit subsidies (relative to other subsidies in the U.S. economy) are too small to merit the attention of most residents in a metropolitan area. Facing little resistance from the public, transit advocates aggressively explore alternative avenues
to fund a new system or extend an existing one. (Winston and Maheshri 2006, 19)

Downs highlights the influence of downtown business interests:

[O]ne way to strengthen the market for office and other space within a business center is to build more off-road transit facilities to serve it. That is undoubtedly why downtown business interests so strongly support construction of new fixed-rail transit systems—especially if they can obtain federal subsidies to cover much of the costs. (Downs 1992, 89)

Winston takes a broader view:

[The] simplest answer to why government repeatedly pursues inefficient transportation policies is that policymakers—appropriately—respond more to political forces than to market forces. Transportation policy has thus become a giant grab bag whose benefits are available to various vested interests—some just get more than others—at the expense of a more efficient transportation system that could save the public at large billions of dollars. (Winston, Summer 1999, 45)

After conducting more than 200 interviews with transportation decision makers, Richmond (2005) concludes that most succumb to the power of myth and delusion and enter the evaluation process already devoted to building rail. They frequently romanticize rail travel, and go through an evaluation process that serves little more than a “ritual function,” one that “gives an aura of respectability to decisions which have been reached on other grounds” (147). Officials focus on forecasting findings that support rail and dismiss those that do not. Just as others speculate that the public overestimates rail’s positive effects (Winston and Maheshri 2006), Richmond (2005, 252) observes a similar phenomenon among public officials. He contends that officials hold beliefs based on incomplete assessments, which causes them to, for example, assume rail

---

10 Lave (1998) notes that observer bias prompts transit riders to assume that load factors are higher than they really are. After all, few people use transit when loads are light.
travel is faster than it really is: “[M]any of those interviewed for this project focused on the rail component of a trip. Other necessary elements of any journey—getting to and from the train stations—were given little or no attention.”

Economists like Hamer (1976) and Kain go further and bring up the possibility of deception. “The tendency of policymakers to ignore the abundant evidence on the superiority of high-performance bus systems is explained by a prior commitment to rail and a willingness to ‘cook the numbers’ until they yield a desired result” (Kain 1999, 384). Kain suggests the use of straw men may be “the most serious flaw of all.”

Nearly all, if not all, assessments of rail transit systems have used costly and poorly designed all-bus alternatives to make the proposed rail systems appear better than they are. In some cases, the use of badly designed alternatives is intentional, while in others a lack of interest in developing better bus systems may account for the inadequacies of the all-bus alternatives. (Kain 1992, 487)

Commenting on a rail proposal in Hawaii, Hamer (1978) offers a similar critique. He charges that consultants recommended rail without first determining whether some kind of express bus network might be a better choice.

Daniel Klein (2005, 19) suggests that rail-transit projects illustrate “the people’s romance,” wherein citizens warm to the idea of a collective effort, endeavor, and experience. In the people’s romance, people imagine that everyone in the community enters into a common sentiment, satisfying a basic yearning for the encompassing coordination of sentiment. Large, visible government projects have this aspect, and we observe that rail promoters often argue that the rail project will help to make the city prominent or “world class.” For economists, however, romantic appeals do not seem to cut much ice, as they almost never count such factors among the benefits of a rail project. Garrett (2004) is somewhat of an exception.

The overall cost of light rail, and the cost of providing rail transit to the poor, can certainly be justified if society obtains some intangible benefit (e.g. pride, generosity, compassion) from knowing that light rail exists in a community. This is similar to the community pride argument made in favor of using tax dollars to finance the
construction of professional sports stadiums. Measuring these intangible benefits that society may receive, however, is difficult. So, although providing light-rail transportation is very costly, each community must weigh the cost with the tangible and intangible benefits it receives from light rail. If these benefits are high enough, then the dollar cost can certainly be justified. (Garrett 2004, 13)

However, counting such intangible factors would open the door to hairy questions about cultural benefits and costs—Klein, for example, argues that the people’s romance is something humankind should not gratify but rather learn to subdue, like the penchant for sweets.

EXAMINING THE EXCEPTIONS

In some cases, it is easy to identify enthusiastic supporters of rail transit. Haynes Goddard, Khalid Bekka, and David Lewis seem to be clear-cut examples of economists who see hearty benefits from rail transit. Lewis contends that light rail investments often yield benefits that exceed costs.

No economic analysis can claim to quantify every conceivable category of the benefit and cost arising from transit investments. Yet, when the major effects are taken into account, light rail investments can be seen to generate economic benefits that exceed the costs of their construction and operation. . . . This is not to say that every light rail proposal is an economic winner; each project proposal must be examined on its merits. (Lewis 2006, 3)

Shapiro, Hassett, and Frank (2002) do not distinguish among modes, but they see great benefits from public transit in general.

---

11 At the time this paper was published David Lewis was President of HLB Decision Economics, an architectural-engineering-consulting firm.
Greater reliance on more fuel-efficient means of travel, especially use of public transportation is the key to the United States achieving greater energy independence and environmental progress. . . . Both pragmatism and patriotism can become catalysts for much greater use of public transportation. As a practical matter, increasing transit may be one of the most feasible—and desirable—strategies for sharply reducing our dependence on foreign oil and making historic strides in environmental quality. As an act of civic commitment, many Americans may view riding public transit, even on a limited basis, as a small but important contribution to our country’s well being. (Shapiro, Hassett, and Frank 2002, 33)

However, in many cases it is rather difficult to discern the degree to which the “exceptions” are exceptional. As was discussed above, Zaretsky (2004, 1-2) writes that “compared with single-occupant automobiles, public transportation, especially rail systems, is a much more efficient way to move people around a metropolitan area.” He speaks favorably of rail transit’s ability to reduce energy use, pollution, and traffic congestion. Yet Zaretsky also brings up concerns that are common to other economists, such as the relatively low use of rail transit and how the rise of suburbanization makes it more difficult for rail transit to succeed.

Zaretsky also worries about rail’s recovery ratio. “One problem light rail faces is that passenger revenue is never sufficient to cover operating costs. . . . [A]pproximately two-thirds of a light rail’s operating expense must be either raised through local taxes or subsidized by the government.” Zaretsky wonders about the feasibility of rail but then seems to reassert its value.

As economic theory tells us, a firm should generally not continue to operate if it cannot generate enough revenue to cover its operating costs. In some sense, then, government is operating a system that should shut down. Of course, one response to this argument is that the rail system’s direct operating costs are not adjusted for the indirect savings from lower energy consumption and the lessening of pollution and congestion. In addition, public transit contributes a positive byproduct by providing affordable transportation for those who cannot afford
other means of transport, like owning a car or taking a taxi. Although these adjustments to operating expenses are not directly measurable, they can play an important role in determining the viability of a system. (Zaretsky 1994, 8).

Zaretsky cites Washington, D.C.’s rail system, which he notes “is considered relatively successful.” Likewise, Voith (2005) also views that system in a favorable light.

Baum-Snow and Kahn cite the fact that new rail transit systems tend to provide service to fewer people than existing systems because the density of the areas served are lower. They argue that this is evidence for the declining marginal productivity of additional rail transit lines. The Washington experience, however, suggests that there are positive network effects associated with extensive development of quality rail facilities. In the Washington area, rail investments have seen extensive development in areas served by the system, increasing both ridership and density. (Voith 2005, 55)

Voith further highlights San Francisco and Boston as areas where rail has made transit relatively successful. As noted above, Voith also sees small positive land use effects of rail transit and he does not view the large proportion of bus-to-rail substitutions as a negative outcome. Perhaps, he argues, bus riders would not have remained transit users if not for investments in rail. Again, given his relative supportiveness of rail transit, it is probably significant that Voith writes that economists’ “dominant view” is that “rail transit investments generally have been ineffective and expensive, and the benefits do not justify the costs” (Voith 2005, 52). He continues:

While I share the view that some rail transit investments have not had large, positive impacts on their communities, measuring the success or failure of rail transit investments is an endeavor fraught with many challenges, from the definition of the counterfactual, to controlling for other local policy choices affecting the outcome of the investment. In general, I believe that the focus of these analyses should not be based on a particular mode, but
rather on the characteristics of the service as perceived by the transportation consumer. Finally, I would conjecture that evaluating the impact of very large transit investments like the Washington Metro by using marginal analysis is nearly impossible. The high levels of use have resulted in a city that otherwise could not evolve in a similar manner. (Voith 2005, 58)

Nelson et al. also highlight the Washington experience. They find that D.C.’s rail (and bus) system offer substantial net benefits, but advise against using this example to make general conclusions.

An obvious limitation of this study is that results reflect specific features of the Washington metropolitan area, including the geography of income distribution, relative importance of public transit, level of carpooling, degree of utilization of HOV lanes, and the fixed central economic activity of the federal government. As Baum-Snow and Kahn (2000) argue, the D.C. region is one of the most promising settings for a major transit system in the country. The large benefits found here should not be taken as evidence in support of transit investment in dissimilar locations. (Nelson et al. 2006, 2)

And since “most of the costs are sunk, and there is little prospect that the system will be disassembled,” Nelson et al. consider estimating the total benefits of the existing transit system to be “an interesting academic and political question,” but “something of a moot point practically.” What they find more significant is “whether the current scale of service is near optimum” (18). The researchers do not advocate large-scale expansions.

Although the current level of investment in transit in the Washington area is not optimal, it is reasonably close. Furthermore, although the value of the transit system as a whole is unquestionable, the net gains from moving the baseline to the optimum (assuming no other concurrent instruments, like road pricing) are trivial compared with the net benefits of the system. Similarly, moving from the optimum to a point of lower provision results in trivial losses. (Nelson et al. 2006, 21)
CONCLUSION

The public debate often strays from some simple considerations that are fundamental to economic analysis. It is not enough to show that rail (or any other public policy) provides benefits. Benefits must outweigh costs. And if it is to be viewed as the best choice, policymakers must be able to determine that it is more effective than all other options at achieving a particular goal. Yet most of the public debate, indeed much research, gives short shrift to considerations of opportunity cost.

The debate is further complicated by the general lack of agreement regarding which goals rail transit should even attempt to achieve. We have aimed to explore most of the core issues that are important to economists, but we have not examined all justifications for rail transit. Policymakers might consider whether it is reasonable to expect rail transit to make good on a wide array of goals and whether insisting on certain goals (e.g., getting comparatively well-off motorists out of their cars) could make it more difficult to achieve other goals (e.g., serving the transit dependent). Further, the success of rail transit will, to a large degree, be determined by trends, demographics, policies, and other factors that vary greatly from one community to another. If a community exhibits the right mix of features, rail can indeed be successful.

Economists also note the importance of the nature of public management in the provision of transportation services. Rail is affected by (mostly) the same incentives that affect other transit modes and it can be difficult to determine the degree to which rail’s performance is the result of public management or the features inherent to the mode. A liberalized transportation regime (especially one that embraced pricing) would likely look much different than what we have now, and economists are often enthusiastic about what kinds of services might emerge and how transit might expand market share.

What conclusions can we draw from economists’ view of rail transit? Economists demonstrate a range of views regardless of the particular justification examined. Still some justifications receive more support than others. Although the matter is filled with disagreement and caveats, economists appear to be the most optimistic about rail transit’s impact on economic development, especially its impact on residential housing values. Economists seem to be less optimistic about rail’s ability to achieve environmental improvement and serve the transit dependent poor. Economists seem quite pessimistic about rail’s ability to achieve key transportation-related goals, like
reducing congestion, and they tend to see other modes, primarily bus, as more functional and worthwhile. Economists often attribute rail’s political success to their belief that decision-makers are motivated by rent-seeking and romantic factors. Of those economists who offer a big-picture view, there appears to be wide, though not unanimous, agreement that rail’s costs exceed its benefits. And it seems that almost all economists who write about rail agree that various demographic features, such as suburbanization, the declining influence of central business districts, and increasing wealth will make it increasingly difficult to design successful rail systems.

REFERENCES


Goddard, Haynes C. 2004b. Comments on Castelazo and Thomas A. Garret Light Rail Boon or Boondoggle. Posted on lighttrainnow.org August 10.


Lave, Charles. 1994. It Wasn’t Supposed to Turn Out Like This: Federal Subsidies and Declining Transit Productivity. *Access*, No.5, Fall.


Transportation Authority conference “Re-inventing Transit: Move into the Millennium,” Chicago.


ABOUT THE AUTHORS

Ted Balaker is the Jacobs Fellow at Reason Foundation. His research and writings focus on urban policy, globalization, and workplace issues. Previously, Balaker worked at ABC Network News, producing segments and documentaries on government reform, regulation, the environment, transportation policy, and other issues. Balaker graduated Phi Beta Kappa from the University of California, Irvine with bachelor’s degrees in Political Science and English. He is co-author (with Sam Staley) of The Road More Traveled: Why the Congestion Crisis Matters More Than You Think, and What We Can Do About It (Rowman & Littlefield, 2006). Email: ted.balaker@reason.org.

Cecilia Joung Kim is originally from Daejeon, Korea but spent her childhood in York County, Virginia. She subsequently moved back to Korea to learn her native language and culture. She received a Bachelor and Master of Urban Engineering from Hong Ik University in Seoul and recently a Master of Planning at the University of Southern California. Her concentration is in urban design, though her interests cover all areas of planning from transportation to economic development. Email: joungkim@usc.edu.
Editors,

I was interested in the rather tense exchange on money in colonial America in the January issue of this journal.

To master their primary material, historians have no choice but to specialize in time and place. When it comes to interpreting their material, however, this specialization can be a handicap. Economic behavior tends to find much the same solutions to the same problems. So when practice in a particular case is obscure, light may be shed by practice in similar circumstances in other times and places.

I have studied money in pre-industrial Europe where a similar monetary chaos was endemic (http://www.dartmouth.edu/~mkohn/, Ch. 7). People in Europe found various ways to adapt to this situation. First, they denominated contractual payments in a money of account that was of relatively stable value. Second, they did their best to avoid the need for settlement by offsetting one debt against another, sometimes assigning third-party debt for this purpose. Third, when settlement was unavoidable, they did not generally settle in the money of account. Indeed, in some cases, as with the Florentine lira affiorino, this was impossible, since the money of account was entirely imaginary. The lira affiorino had no physical existence, but its value was defined as 20/29 of a florin—an actual gold coin. Settlement was made in whatever coin was available, valued at a varying—often negotiated—rates of exchange against the money of account. For example, a debt of 8 lira affiorino might be settled with a combination of 70 grossi (a large silver coin valued at, say 14 to the florin) and 116 quattrini (a copper-alloy coin valued at, say 16, to the grosso).

Moreover, as it happens, I have recently been studying China. There too, in the 12th through 14th centuries, monetary chaos ruled (von Glahn, 1996).
Not surprisingly, people responded in much the same way. They denominated debts and prices in a money of account—copper cash—the traditional form of money. However, copper cash was no longer generally available as a means of settlement. Instead, people settled amounts denominated in cash in varying combinations of paper money, unminted silver, silk cloth, rice, and whatever else was agreeable to the parties.

It seems to me plausible that practice must have been much the same in the colonies.

Meir Kohn  
Dartmouth College  
mkohn@dartmouth.edu

Reference  