Econ Journal Watch

Scholarly Comments on Academic Economics

Volume 3, Number 2, May 2006

COMMENTS

ECONOMICS IN PRACTICE

DO ECONOMISTS REACH A CONCLUSION?

CHARACTER ISSUES

INTELLECTUAL TYRANNY OF THE STATUS QUO

Econ Journal Watch,
Volume 3, Number 2,
May 2006, pp 210-235.

Is the Swedish Welfare State A Free Lunch?

ANDREAS BERGH*

A COMMENT ON: PETER H. LINDERT. 2004. *GROWING PUBLIC: SOCIAL SPENDING AND ECONOMIC GROWTH SINCE THE EIGHTEENTH CENTURY*. CAMBRIDGE UNIVERSITY PRESS.

Abstract

WITH HIGH TAXES, HIGH PUBLIC EXPENDITURES, AND A SOCIAL policy many would call progressive, Sweden is a special country for social scientists. Peter H. Lindert makes special use of Sweden in his book *Growing Public*, first published by Cambridge University Press in 2004 and translated into Swedish in 2005.[1]

The main part of the book aims to describe and explain the expansion of western welfare states.[2] Lindert uses Sweden to explain why the welfare state appears to be a free lunch, what he calls the "Free-Lunch Puzzle." I argue that Lindert misrepresents Sweden when it comes to work incentives for the poor, employment of women, and employment of the elderly, and that he does not pay sufficient attention to the many reforms undertaken in Sweden since the late 1980s.

* The Ratio Institute, Stockholm, Sweden SE-102 42.
The author would like to thank Niclas Berggren for valuable comments.
[1] Lindert (2004), Lindert (2005).
[2] For comments and critique of this part, see for example Margo (2004).

THE SO-CALLED FREE-LUNCH PUZZLE

The book's marketing material and an earlier working paper (Lindert 2003) suggested that welfare states have been a free lunch. The free-lunch puzzle, however, refers to a less controversial statement: Lindert has not been able to find a robust and statistically significant negative effect of "social transfers" on economic growth. Lindert's definition of social transfers is very specific. It includes only the following (Lindert 2004, 6-7):

- Poverty relief ("welfare" in America)
- Unemployment compensation
- Non-contributory pensions
- Public health expenditures
- Housing subsidies

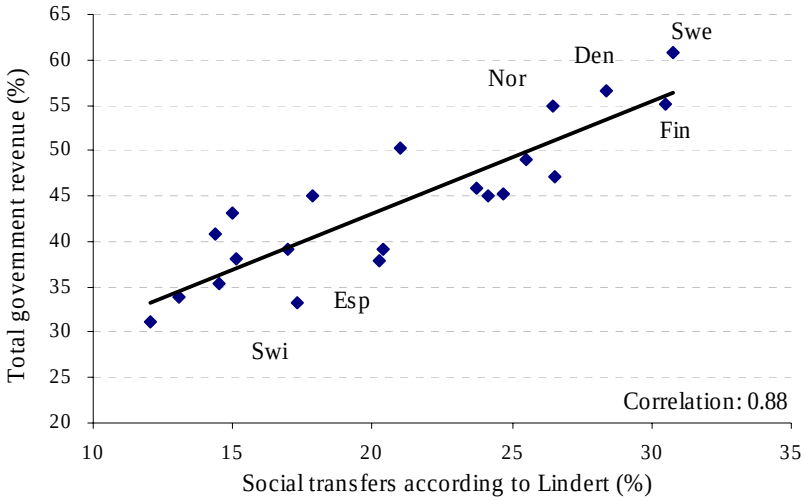
More details are provided by the datasets available at the publisher's website (Link). Public health expenditures include sickness benefits. Housing subsidies do not include public housing. Unemployment compensation does not include the costs of labor market programs. Public subsidies to corporations are not included.

Lindert acknowledges that other studies find a negative effect of different measures of government size on growth. [3] In fact, Lindert stresses that these other studies typically use total taxes or expenditures as a share of GDP rather than social transfers as defined by Lindert.

Using Lindert's data, it is easy to examine the correlation between Lindert's social transfers and public expenditure as defined by the OECD. For the countries used in Lindert's study, the correlation was 0.82 for the year 1995. The correlation between social transfers and total tax revenue was 0.88 (shown in Figure 1). Compared to using measures based on total taxes or total expenditures, using the measure of social transfers chosen by Lindert makes Sweden look slightly less extreme.

[3] For example Persson and Tabellini (1994), Commander et al. (1997), and Fölster and Henrekson (1999).

Figure 1: The correlation between Lindert's social transfers and OECD's total government revenue (percent of GDP, 1995)



Source: Lindert (2004) and OECD.

I refrain from delving into the econometric evidence of the free-lunch puzzle because it has long been known that it is hard to produce robust results regarding the effect of government size on growth[4], and my theoretical explanation of this is no other than Lindert's: 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.

For this reason, the term "Free-lunch puzzle" is highly inappropriate: The murkiness of the relationship between "social transfers" and economic growth is not a puzzle, and the welfare state is not a free lunch. Lindert's regressions seem to leave open the possibility of constructing a welfare state in a way that does not harm growth. But Lindert makes the specific

[4] For example, Fölster and Henrekson (1999) and Agell et al. (1999) obtain different results using the same datasets. The debate has continued—see Fölster and Henrekson (2006) and Agell et al. (2006). See also Gordon and Wang (2004).

assertion that Sweden has enjoyed its welfare state for "free" (that is, with no diminution of economic growth). In developing the example of Sweden to illustrate his vision, Lindert makes faulty descriptions of the Swedish experience, both in terms of policies and outcomes. A more accurate description of Sweden suggests that Sweden's welfare state is associated with substantial costs.

CORRECTING THE DESCRIPTION OF SWEDEN

Lindert paints a somewhat naïve picture of Sweden, and some of his explanations of the free-lunch puzzle simply do not work in the case of Sweden.

Work incentives for the poor in Sweden

In the chapter "Keys to the Free-Lunch Puzzle", Lindert opens with a critique of means-tested social welfare. With means-testing, as income increases, benefits are reduced. Such an arrangement weakens work-incentives for the poor. Lindert claims that the means-tested policy in the United States during the Reagan years "discouraged work more than in later years or in the *true welfare states*" (230, italics added). By "the true welfare states," Lindert probably means the welfare states usually labeled universal or encompassing, i.e. the Nordic welfare states.

Later in the chapter, Lindert writes that "the poor may face lower work disincentives in the welfare state" (245), and suggests that this might be one reason for the absence of a negative effect of social transfers on growth. To illustrate work incentives, he displays marginal net tax rates for a lone parent with two children in the United States and the United Kingdom, and compares these with "characteristic" Swedish rates.[5] Finally, Lindert suggests: "at the bottom of the income spectrum . . . the universalist welfare states may well have lower marginal net tax rates than the lower-budget countries, which emphasize strict means testing" (249).

[5] The marginal net tax rate measures the extent to which an increase in personal income is offset by taxes and reduced benefits.

For a Swedish social scientist, this suggestion is remarkable. The system used for poverty relief in Sweden (social assistance, in Swedish *socialbidrag*) is a typical example of a targeted, means-tested system, creating poverty traps because benefits are reduced crown-for-crown against work incomes. Lindert, however, makes no mention of the social assistance system. In Bergh (2004) I show that the social assistance system is indeed a prominent example of a non-universalistic sub-system in the universal welfare state of Sweden.

To be eligible for social assistance, the individual must show no assets. The benefit is paid in cash. One part consists of general cash benefits, given to cover expenditures on food, shoes, clothes, leisure activities, phone, newspaper and TV. For 2006, the amount of the general benefit is 3420 SEK per month (approx. US $430) for a single adult with no children. Another part covers the individual's actual expenditures on housing, electricity, travel, home insurance, union membership, and fees to unemployment insurance. Also, in some cases expenditures on furniture and dental and medical care are also covered. Because rent and electricity benefits alone can easily sum to 4000 SEK for a single adult, the total benefit level is very high by international standards: In the calculations by Carone et al. (2004), a single adult is eligible for social assistance if her income is below 49 percent of the earnings of an average production worker.[6] When the specific expenditure posts covered by social assistance is higher, for example if rent, electricity and travel costs are high, the part that is not tied to specific consumption is relatively smaller, and it is typically less than 50 percent.

In Lindert's table 10.2, labeled "Hurdles in the path out of poverty?" he reports marginal net tax rates "between 30% and 50%" for Sweden. According to Lindert, the Swedish rates shown are "the averages of those generally characteristic of a single adult student, a couple with children in day care, and an absent parent subject to child support, all in Stockholm 1991" (246). If the groups identified do receive social assistance, the numbers are wrong. More likely, the groups do not receive social assistance and the numbers are correct. In this case, however, the chosen groups are hardly suitable for Lindert's purpose, which is to compare work incentives for the poor in Sweden, the United Kingdom and the United States.

[6] Given that the full-time earnings of an average production worker in Sweden was 230 000 SEK in 2001, this estimate seems to be a little on the high side. (1 US$ is approximately 8 SEK)

Students and couples with children in day care are typically not poor in Sweden.

Those who are poor in Sweden typically face very low work incentives. In fact, low work incentive for the poor has been an issue on the academic and political agenda in Sweden for a long time—see for example Lindbeck (1997), Söderström et al. (1999), and Ministry of Finance (2003). Different measures of work incentives that include effects of benefit reductions are also a central part of the EU's Lisbon strategy, and are closely monitored by governments.

Table 1 below shows data from 2001 in the Swedish employment strategy. The numbers indicate the income-gain portion that is offset by taxes and reduced benefits for an individual moving from being unemployed to employment (and hence losing some or all social assistance). The columns show four cases: when the person earns a third, half, two thirds and 100 percent of what an average production worker earns. We see that the marginal net tax rates are very high for people with low wages and no income-related unemployment benefit.

Table 1:
Marginal net tax rate for an individual moving from being unemployed and receiving social assistance to employment (and hence losing some or all social assistance)

	Wage, percent of APW[7]			
	33	50	67	100
Marginal net tax rates of person with income-related unemployment benefit	59	86	87	85
Marginal net tax rates of person without income-related unemployment benefit	100	98	82	67

Source: Link.

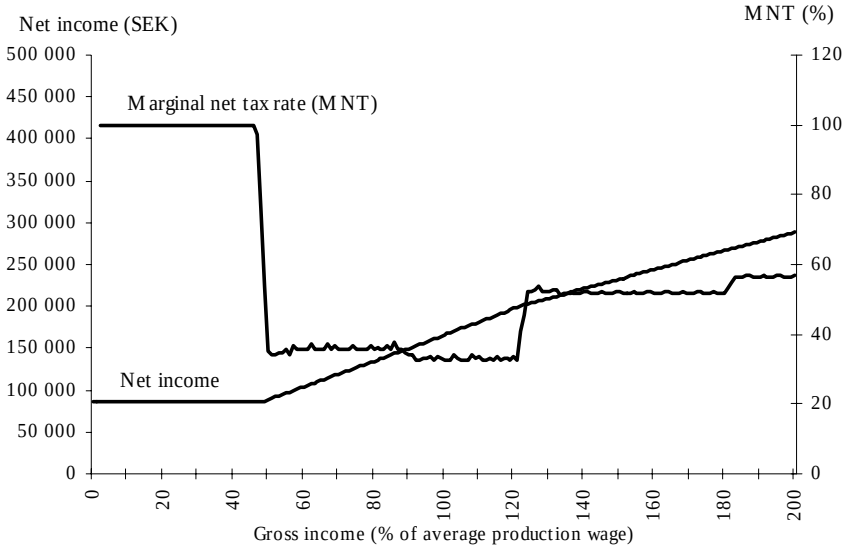
It should be noted that the weak work incentives affect a significant portion of the Swedish population. The Swedish social assistance is not a marginal phenomenon: During the nineties, the share of households that received at least some social assistance was between 8 and 11 percent every year (Socialdepartementet 2000).[8] And although the standard measure of poverty has only five percent of the households below 50 percent of the

[7] APW is the full-time earnings of an average production worker.
[8] Among single women with children, take up rate was 29 percent in 1999.

median household income, 12 percent of the population are found below the income levels indicating eligibility for social assistance.[9]

The poverty trap in Sweden arises because social assistance is completely reduced against all individual earnings. To illustrate work incentives for different incomes, it is useful to plot net-income curves against individual gross income. Figure 2 shows the situation in 2002 in Sweden for a single person without children and not eligible for unemployment insurance. The net-income curve incorporates social assistance, other benefits, and income taxes, and hence represents the individual's budget set. When the net income curve has a slope of zero, the marginal net tax rate is 100 percent, indicating that an increase in gross income for the individual has no effect on her net income—a complete poverty trap. We see clearly that material work incentives for poor people are close to non-existent in Sweden.

Figure 2: Work incentives in Sweden in 2002 for a single person with no unemployment insurance

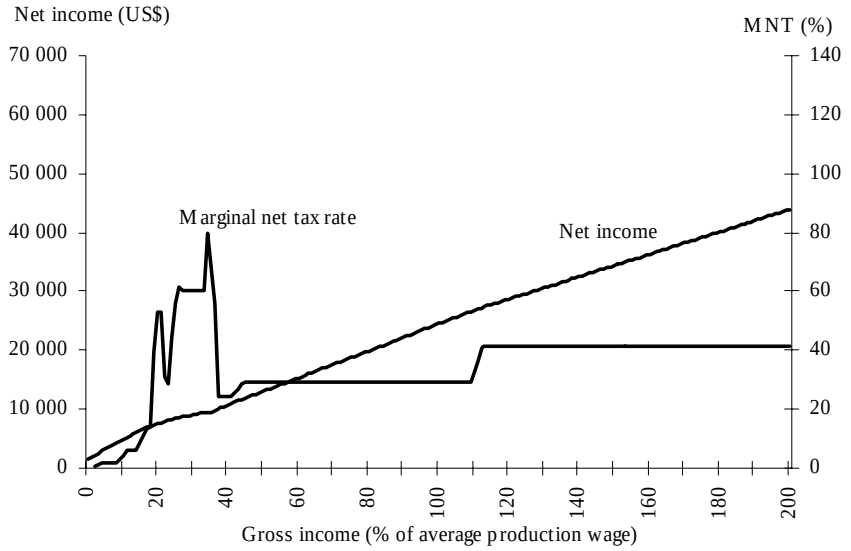


Source: OECD (2004)

[9] Ministry of Finance (2003).

Figure 3 shows the situation for the same type of individual in the US. Whereas the material work incentives for the poor in Sweden are non-existent, in the US they are rather strong.

Figure 3: Work incentives in United States in 2002 for a single person with no unemployment insurance



Source: OECD (2004)

Clearly, the difference between Sweden and the United States is fundamental, and one might suspect that the OECD-data for example does not include state level benefits. But the documentation clearly state that because tax and benefit systems vary from state to state, OECD has used the State of Michigan to represent a typical manufacturing region. OECD states that Michigan benefits are somewhat above the average for all States.[10] As for Sweden, Figure 2 based on OECD-data looks close to identical to the one I myself produced for the 1995 situation in Sweden (see figure 5 in Bergh (2004)). I have also verified that the qualitative and fundamental differences between the United States and Sweden is not

[10] All documentation and data for OECD (2004) is available online: Link.

driven only because the comparison is made for a single earner with no children—the same point could be made with any household type.

The lack of material incentives in Sweden is to some extent made up for by rules requiring social-assistance recipients to be actively looking for jobs. Also, in some cases, they are required to participate in public employment projects. Nevertheless, by comparing Figures 2 and 3, it is evident why in the United States the poverty debate is focused around the 'working poor,'[11] whereas the Swedish debate is centered around benefit dependency and separation from the regular labor market.

The level and length of the flat segment on the net income curve depends on the type of household and the amount of approved expenditures covered by social assistance. More children and higher approved expenses mean higher total benefit and thus also a bigger poverty trap, in the sense of a longer flat segment on the net income curve.[12]

The difference between work incentives in Sweden and the United States can be explained by fundamental differences in the tax and benefit systems. Swedish policymakers have opted for high social assistance as described above *and* high taxation of low-income earners. American policymakers have chosen the opposite strategy: The Earned Income Tax Credit essentially subsidizes low wage employment, while those who do not work receive very little public support.

The taxation of the very poorest in Sweden is higher than even many Swedish social scientists probably realize. Recipients of social assistance pay income taxes at roughly 30 percent on their earnings. When a social-assistance recipient increases her income by 100 SEK, she pays 30 SEK in local income taxes, and social assistance is decreased by 70 SEK. Moreover, employers pay wage taxes at roughly 30 percent of those wages. Thus, when the social-assistance recipient increases her income by 100 SEK, the public budget is improved by approximately 130 SEK, but her private budget is improved by 0 SEK.

To sum up, the Swedish system sends a clear signal: Work only pays materially if you can get a job that pays at least half the earnings of an average production worker. The system in the United States is a stark contrast: Even if the only job you can get pays less than a fifth of that of the average production worker, you are better off working.

---

[11] See for example Ehrenreich (2001).

[12] Note also that because the cash amount is fixed, and the rest of social assistance covers expenditure, there are no economic incentives for individuals to cut down on approved expenditure.

Table 1 showed that weak work incentive in Sweden is not only a problem for the poorest. High-level, income-related social insurance benefits imply that the material improvement when accepting a job is small also for many middle-income earners. However, because of the construction of social assistance, the complete poverty trap with 100 percent marginal net tax rate applies only to the poorest.

In his table 10.2, Lindert, for some reason, reports marginal net tax rates for the United States and the United Kingdom both with and without the Earned Income Tax Credit (EITC), and the similar Working Families Tax Credit (WTFC) in the United Kingdom. His arguments hold in theory when comparing universal welfare states to those relying on strict means testing. But they do not apply to the Sweden-United States comparison. As clearly shown by Lindert's data, the EITC has had a very big positive impact on work incentives for the poor, whereas Sweden has nothing similar to alleviate the negative incentive effects of its strictly means-tested social assistance scheme.

Lindert asks: "If welfare states really have lower marginal tax rates at the top and bottom of the income spectrum, but higher tax rates in the middle, do they discourage work more or less than the low-budget governments of Japan, Switzerland and the United States" (249)?

The question is odd because the structure of the marginal net tax rates in Sweden is exactly the opposite of Lindert's description—high for low and high-income earners, and lower for middle income earners (see Bergh (2004)).

I can only conclude that Lindert's description of work incentives in Sweden is wrong. For this reason, one of Lindert's keys to understanding the free-lunch puzzle is fundamentally shaky. On page 245, a sub-heading reads "The Poor May Face Lower Work Disincentives in the Welfare State." Well, in theory they might, but in Sweden, they don't.

Employment of women in Sweden

A second point where Lindert misrepresents the situation in Sweden is employment, especially for women. He writes, "relative to other OECD countries, Sweden's institutions seem to produce greater employment, especially jobs held by women and the elderly, with positive effects on GDP" (281). And, "virtually all of Sweden's employment growth between the 1960s and the early 1990s consisted of jobs for women" (282).

From these quotes, it is evident that Lindert wants to use high female employment as another explanation of the free-lunch puzzle. To support his claim he uses OECD data on employment ratios (Table 11.3, 282). The employment ratio, however, is a very blunt measure of employment. It tells you whether a person is employed, but not how much she works. High and increasing employment among women or elderly can explain the Swedish GDP level and growth only if Sweden scores high when it comes to actual number of hours worked, as opposed to employment ratios.

To explain high female employment in Sweden, Lindert combines two facts: "Sweden's women get better pay relative to men than in any other OECD country" and "women have a more elastic labor supply than men" (282). He also stays with his theory that high gender equality is an explanation of the free-lunch puzzle: "Women have a more elastic labor supply than men, and that converts gender gaps into effects on *total work* and GDP" (287, italics added). When we consider the number of hours actually worked instead of employment ratios, however, total female employment in Sweden is actually lower, not higher, than in the United States.

It has long been known that the seemingly high gender equality on the Swedish labor market is much less impressive once you look at the actual numbers of hours worked. In a paper titled "Women and Market Work: The Misleading Tale of Participation Rates in International Comparisons," Jonung and Persson (1993) showed that behind a gender difference of only 5 percentage points in labor force participation in 1998, there is a difference of 12 percentage points when the actual number of hours worked are taken into account. By using employment ratios, Lindert hides the fact that Swedish women work fewer market hours than Swedish men, and they spend more time in unpaid non-market work.

By taking into account the actual number of hours worked, Henrekson (1998) reports adjusted employment ratios comparable to those reported by Lindert. The results for Sweden and the US are shown in Table 2. When actual working hours are taken into account, we find that Lindert is mistaken about Sweden having higher female employment.

Table 2:
Female employment and hours worked in Sweden and the US (1994)

	Employment ratio[13]	Hours worked per week	Employment ratio adjusted for hours worked[14]
USA	66 %	36	62 %
Sweden	71 %	28	51 %

Source: Henrekson (1998).

International comparisons of working hours are difficult because of variation in definitions and methods. As shown in Table 3, alternative sources differ in their reported levels. But they still confirm that the big difference between the United States and Sweden is that both men and women in the United States spend more time in market work compared to their Swedish counterparts.

Table 3:
Market work per week for men and women
NB: Alternative sources[15]

Market work per week (avg hours)		
	USA	Sweden
Men	42	38
Women	36	33

I am inclined to reject Lindert's elastic labor supply explanation of Sweden's high female employment ratio. What is the better explanation? Again, the answer seems obvious and not very controversial: The expansion of the public sector. Lindert tells us that "all of Sweden's employment growth between the 1960s and the early 1990s consisted of jobs for women" (282). He does not tell us that in Sweden, there has been no net

[13] These numbers are reported by Lindert in his table 11.3.
[14] This equals number of hours worked per week * (52/2000)*employment ratio.
[15] Source: Table 17 in Nelander and Goding (2005) for Sweden (data for 1994), and table 1 in Rones et al. (1997) reporting data from 1995. Note also that according to standard OECD data, Americans (men and women together) spend approximately 200 hours more in market work every year compared to Swedes, see for example OECD Productivity Database, available at online: Link..

growth of jobs in the private sector since 1950, as shown in Figure 4, taken from Davidsson and Henrekson (2002).

Figure 4:
Cumulative change of private employment, government employment and population in Sweden (in thousands), 1950–98

Source: Davidsson and Henrekson (2002).

Clearly, there is important information contained in both the employment ratios and the number of hours worked. By focusing on the 5 percent that separates Sweden from the United States in terms of employment ratios, one could argue that for these women, having an independent income from the public sector is preferable to being

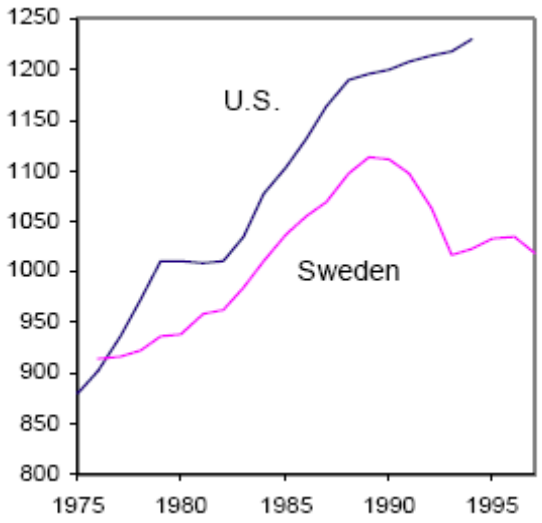
dependent on a husband. However, when Lindert is aiming to explain "total work and GDP," displaying only employment ratios is not convincing.

Furthermore, by ignoring the distinction between employment ratios and hours worked, Lindert's comparisons between Sweden and the United States become very biased. Hakim (1996) has shown that when it comes to female employment, Sweden is surprisingly similar to other countries: "[t]he rise in female employments rates is due primarily to the creation of a new part-time workforce (the USA being a rare exception)" (61). Highlighting American exceptionalism, Hakim adds "Only the USA has had a steady and accelerating growth in women's employment" (65).

Hakim's reading of the differences between Sweden and the US is robust to the choice of years for comparison.

Figure 5 depicts the development of annual hours worked over time in Sweden and the United States. From this development it seems to be clear that high female employment is indeed a factor explaining growth differences between Sweden and the United States. But the factor favors the United States, not Sweden.

Figure 5: Average Annual Hours Worked Among Working Age Women in the U.S. (1975–94) and Sweden (1976–97)



Source: Henrekson and Dreber (2005).

Because Lindert fails to identify these fundamental differences between Sweden and the United States, he also fails to discuss possible theoretical explanations of the pattern. Once again, the answer is obvious and uncontroversial: To finance the expansion of the public sector (depicted in Figure 4), Sweden has increased labor taxation substantially.

High tax wedges on labor have several effects. In this context, one important effect is that it becomes more expensive for a person to hire another person to do any type of job. When one person hires another person, labor taxation enters repeatedly in the total tax wedge. First the hiring person must earn money and pay taxes. With her after-tax income, she hires another person, who once again has to pay taxes. In addition to income taxes, the same goes for wage taxes if such exist.

DuRietz (2004) computes the total difference between what the hiring person must earn and what the person she hires will keep in a number of countries, assuming that the hiring person earns twice as much as an average production worker and the hired person earns the average production-worker wage. Based on his data, I have computed the total "person-to-person" tax wedge for a number of countries, in a way that makes them comparable to standard business-to-person tax wedges.[16]

Table 4 shows that this person-to-person tax wedge is much higher in the Nordic welfare states compared to the US. It should be noted that these tax wedges are extreme also by the standards of high tax countries. For comparison, I have also included standard (business to person) OECD tax wedges in table 4. We see that the tax wedge when a business hires a person is much smaller.

[16] For example, DuRietz shows that to hire a person in Sweden that will keep 1000 SEK, you must earn 2611 SEK, which in turn means your employer must pay an additional 6013 SEK. The tax wedge will be (6013-1000)/6013 = 0.83, indicating that 83 percent of the total cost is taxed before 1000 SEK eventually reaches the one who does the job.

Table 4:
Tax wedges and female hours worked in different countries

Country	Person to person tax wedge, (%)	Standard OECD tax wedge (%)	Hours worked per week for women
Belgium	90	48	31.9
Denmark	85	38	31.3
Germany	85	44	33.4
Sweden	83	46	27.6
Italy	81	48	34.4
Netherlands	80	40	22.9
France	75	44	34.9
Ireland	69	32	32.8
UK	68	30	29.5
Portugal	67	30	37.7
Spain	67	36	34.7
USA	60	28	35.8
Japan	52	17	37.1
Luxembourg	-	34	34.4
Greece	-	35	37.6

Source: Person to person tax wedge computed from table 2 in DuRietz (2004), data for 2001. Standard tax wedges are from OECD (2000) Taxing Wages, data for 1995. Female hours worked are from Henrekson (1998), data for 1994.

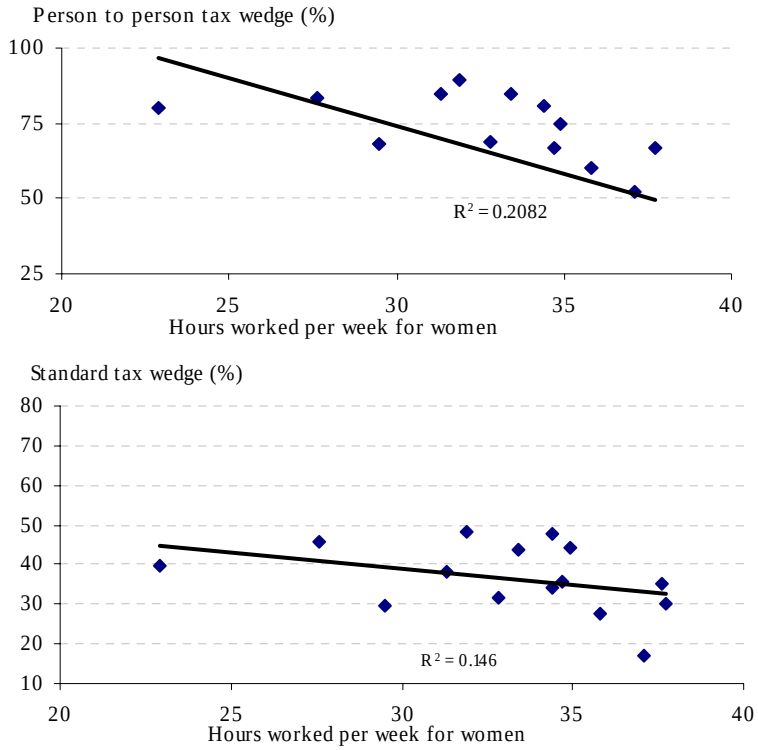
The simple alternative to hiring someone to wash your windows or repair your plumbing is to do it yourself. The higher the tax wedge, the higher must be the productivity difference between the hired professional and doing it yourself. If the tax wedge is too high, you do not work a full day and hire a professional to clean or windows or repair your plumbing, but rather take the day off (foregoing income) and do it yourself. This suggests that one reason for the low hours of female market work is that tax wedges make it harder for working women to hire someone to do household work.

If the tax wedge explanation holds, there should be a negative relation between female employment in terms of hours worked and tax wedges. As shown in Figure 6, this relationship is clearly visible using standard OECD tax wedges and, even more so, when using person-to-person tax wedges. While the sample is too small to do a multivariate

analysis, it seems unlikely that this correlation is driven entirely by omitted variables.

Thus, I suggest the following mechanism as an alternative to Lindert's explanation based on wage equality. Big welfare states have high female employment rates, because the public sector employs (mainly) women to do certain tasks that (mainly) women did before the expansion of the state, namely, caring for children, the convalescent and unwell, and the elderly. Simultaneously, the taxes necessary to pay for the public sector hinder the market for household services. The do-it-yourself incentive created by the welfare state helps explain why working women in big welfare states less often work full-time.[17]

Figure 6:
Average hours worked per week for women and tax wedges[18]



[17] For further evidence, see Hakim (1996), table 3.3.
[18] Source: See table 4.

Employment of the elderly in Sweden

Lindert also wants to use high employment among the elderly as an explanation of the free-lunch puzzle. Again he uses the employment ratio to illustrate Sweden's allegedly high employment among elderly, and his argument is vulnerable to the same criticism as for female employment: He has not verified that the high employment ratio in the age group 55-64 implies that Swedish elderly actually work more hours.

However, Lindert also uses the retirement age to make his point. On page 289 under the heading "Late Retirement" he writes:

> Knowing that Sweden has generally had a low rate of unemployment, one is prepared for the news that Sweden helped that rate look low by removing older workers from the labor force.
>
> Yet the opposite is true of Sweden. Swedish men work to later average retirement ages than men in any other core OECD country except Japan, Norway and Switzerland [in a footnote attached here, Lindert also excludes Iceland, Korea and Luxembourg].
>
> *Swedish women work to later average retirement ages than women anywhere else in the world. Continuing work at advanced ages is one of the many ways in which Sweden achieves higher GDP per capita.* (289, emphasis added)

Despite the allusion to OECD, OECD data does not support Lindert's claim. For both men and women, the retirement age in Sweden is only a few months higher than the OECD average, as shown in Table 5.

Table 5: Average effective age of retirement (in years) versus the official age, 1997-2002

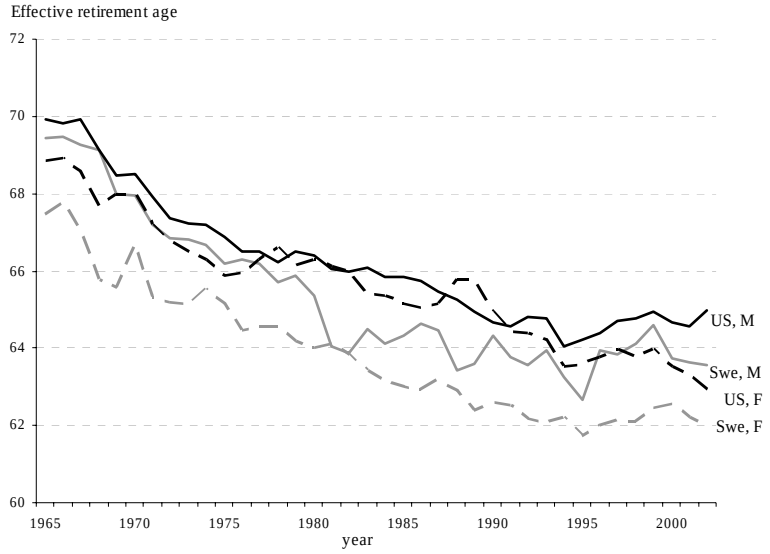
	Men		Women	
	Actual	Official	Actual	Official
Iceland	69.6	67	67.8	67
Mexico	73.8	65	67.2	65
Korea	68.0	60	66.8	60
Ireland	65.2	66	66.2	66
Japan	69.6	60	65.7	60
Portugal	65.8	65	63.5	65
Switzerland	66.6	65	63.2	63
United States	65.0	65	62.9	65
Norway	63.7	67	62.3	67
Denmark	65.3	67	62.1	67
Sweden	**63.5**	**65**	**62.0**	**65**
Turkey	62.5	60	61.9	58
OECD	**63.3**	**63.9**	**61.4**	**62.6**
Canada	63.1	65	61.4	65
Spain	61.6	65	61.3	65
New Zealand	64.3	65	61.3	65
U.K.	63.1	65	61.2	60
Greece	62.4	58	60.9	58
Australia	63.2	65	60.6	62
Italy	61.2	65	60.5	60
Germany	60.9	65	60.2	65
Luxembourg	59.8	65	59.8	65
Finland	60.8	65	59.8	65
France	59.3	60	59.4	60
Netherlands	61.0	65	59.1	65
Austria	59.6	65	58.9	60
Poland	60.9	65	58.8	60
Czech Rep.	62.0	61.2	58.3	59.3
Belgium	58.5	65	56.8	62
Slovak Rep.	59.4	60	56.1	55
Hungary	57.8	62	56.0	58

Source: OECD (2005).

Lindert writes, "Swedish women work to later average retirement ages than women anywhere else in the world", yet women in Iceland, Mexico, Korea, Ireland, Japan, Portugal, Switzerland, United States, Norway and Denmark work longer. This is a remarkable mistake, and it seems appropriate to ponder possible explanations: Perhaps Lindert referred to the official rather than the actual retirement age? Or perhaps he used older data from OECD? Neither of these explanations apply: Official rates are higher in many countries. The 2001 version of OECD's "Society at a Glance" contains data for the period 1983-88 and 1994-99, but those data do not support Lindert's claim.

Figure 7 shows that never in the period of data availability has the actual retirement age been higher in Sweden than in United States. To conclude, Lindert's description of employment among the elderly in Sweden is not supported by data, and thus also fails as a potential explanation of the free-lunch puzzle.

Figure 7:
Trends in average effective age of retirement for men and women in Sweden and USA



Source: OECD (2005).

SWEDEN'S TRAJECTORY: A DIFFERENT INTERPRETATION

It has been shown that Lindert does not succeed in explaining the free-lunch puzzle in general, nor in explaining Sweden's political and economic trajectory. I will conclude by suggesting an alternative interpretation of Sweden's development, one that is not considered by Lindert: Increasing economic freedom.

In a chapter titled "On the Well-known Demise of the Swedish Welfare State," Lindert notes a repeated theme in the Anglo-American press in the period 1977-1998: "That Sweden's economy was falling behind and that its welfare state was to blame" (264). The following quote is typical of Lindert's way of explaining the survival of Sweden's welfare state.

> Given Sweden's insistent egalitarianism, most outside observers also presume that Sweden's tax system is highly progressive, taking over 70 percent of income from rich people at the margin and giving it to the poor. Sweden's actual tax practice is so far from these natural perceptions that we can gain a great deal of insight just by describing features of the system, without trying to quantify its growth effects. (287)

Something upon which Lindert does not reflect is that Sweden's top marginal tax rates peaked at 87 percent in 1979, and remained well above 70 percent throughout the 80s. Thus, for a long time the presumption of "outside observers" was in fact correct. However, perceptions of Sweden have most likely lagged reforms that accelerated in the late 80s.

Lindert notes correctly that during the economic crisis of the 1990s high taxes and public expenditure persisted, and that the degree of universality of the Swedish welfare state did not decrease substantially—see for example Rothstein and Lindbom (2004) and Bergh (2004). But in Lindert's world, the resilience of the Swedish welfare state is used to support his claim that the demise of the Swedish model was mistaken. My alternative interpretation is that many aspects of the Swedish welfare state did in fact cause big problems, and that Sweden's resilience has been the result of a gradual adaptation, in which many of the problems identified by the critics were at least partially addressed.

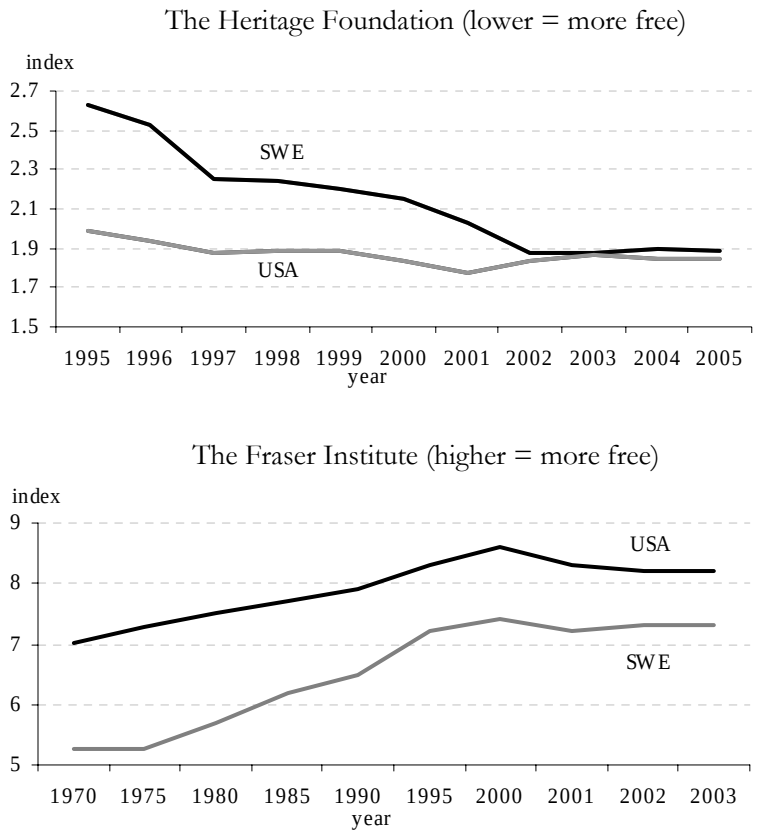
Starting in the 1980s and continuing at least until the mid 1990s, Sweden implemented an impressive series of reforms, increasing the efficiency of its economic system. Table 6 provides a summary.

Table 6:
Some important reforms undertaken in Sweden in the 80s and 90s

1985-89: Deregulation of the credit and currency market.
1990: The initiation of a process of deregulation of markets with a state monopoly: electricity, postal services, telecommunications, railroads and domestic airways.
1990-91: A tax reform decreases the top marginal tax rate from 73 to 51 percent.
1990-97: Central bank independence and a new macro economic policy.
1991-1995: Sweden applies for EU membership and joins the European Union.
1991-2000: A number of state owned enterprises are at least partially sold.
1992: The use of vouchers to introduce competition among schools starts expanding. Vouchers are increasingly also used for child care, elder care and other public services as well.
1993: The legalization of private, for-profit employment agencies.
1994-98: The introduction of a new, partially funded pension system.
1997: The introduction of a new budgetary process with upper expenditure limits.

While these reforms had little effect on quantitative aggregates such as tax or expenditure shares of GDP, they were qualitatively important and contributed to changing the structure of the Swedish economy. One way of illustrating this is to examine the development of economic freedom in Sweden, as measured by the indexes developed by the Heritage Foundation and the Fraser Institute, shown in Figure 8.

Figure 8: Economic freedom indexes in Sweden and USA



Sweden starts from much lower level of economic freedom, but the gap between Sweden and the United States has been closing. Sweden's gains in economic freedom also separate it from other high-tax European countries such as France, Germany and Belgium. The Nordic welfare states all share the fact that they have increased economic freedom faster than the rest of Europe.

To some extent, this development is acknowledged by Lindert. He mentions the tax reform and blames Sweden's poor growth performance on its macro economic policy until the floating exchange rate (unintentionally) adopted in 1992. But that's about it. Lindert does not attempt to describe the major developments indicated by Table 6 and Figure 8. Looking at these changes, the reforms toward increased economic freedom seem to be a good candidate for explaining the survival of the Swedish welfare state.

REFERENCES

Agell, Jonas, T Lindh, and H. Ohlsson. 1999. Growth and the Public Sector: A Critical Review Essay: A reply. *European Journal of Political Economy* 15(2): 359-366.

Agell, Jonas, Henry Ohlsson, and Peter Skogman Thoursie. 2006. Growth effects of government expenditure and taxation in rich countries: A comment. *European Economic Review* 50(1): 211-219.

Bergh, Andreas. 2004. The Universal Welfare State: Theory and the Case of Sweden. *Political Studies* 52: 745-766.

Carone, Giuseppe , Herwig Immervoll, Dominique Paturot, and Aino Salomäki. 2004. Indicators of Unemployment and Low-Wage Traps (Marginal Effective Tax Rates on Employment Incomes). OECD Social, Employment and Migration Working Paper no. 18.

Commander, Simon, Hamid R. Davoodi, and Une J. Lee. 1997. The causes of government and consequences for growth and well-being. World Bank Policy Research Working Paper 1785.

Davidsson, Per and Magnus Henrekson. 2002. Determinants of the Prevalance of Start-ups and High-Growth Firms. *Small Business Economics* 19(2): 81-104.

DuRietz. 2004. Effektivitetsproblem med höga skattekilar. Pp. 165-198 in *Skatter och Värdighet*, edited by N. Karlsson, D. Johansson, and R. Johnsson. Stockholm: Ratio.

Ehrenreich, Barbara. 2001. *Nickel and Dimed.* USA: Metropolitan Book.

Fölster, Stefan and Magnus Henrekson. 1999. Growth and the public sector: A Critique of the Critics. *European Journal of Political Economy* 15(2): 3337-3358.

_____. 2006. Growth effects of government expenditure and taxation in rich countries: A reply. *European Economic Review* 50(1): 219-222.

Gordon, Peter and Lanlan Wang. 2004. Does Economic Performance Correlate with Big Government? *Econ Journal Watch* 1(2): 192-221. Link.

Hakim, Catherine. 1996. *Key Issues in Women's Work*. Great Britain: Athlone Press.

Henrekson, Magnus. 1998. Från en ond cirkel för tjänstesektorn till en utvidgad marknadssektor. *Arbetsmarknad & Arbetsliv* 4.

Henrekson, Magnus and Anna Dreber. 2005. Female Career Success: Institutions, Path Dependence and Psychology. *SSE/EFI Working Paper Series in Economics and Finance No. 574* Stockholm School of Economics.

Jonung, Christina and Inga Persson. 1993. Women and Market Work: The Misleading Tale of Participation Rates in International Comparisons. *Work Employment Society* 7: 259-274.

Lindbeck, Assar. 1997. The Swedish Experiment. *Journal of Economic Literature* 35(3): 1273-1319.

Lindert, Peter H. 2003. "Why the Welfare State Looks Like a Free Lunch." *NBER Working Paper No. 9869*.

_____. 2004. *Growing Public*. Cambridge: Cambridge University Press.

_____. 2005. *Välfärdsstatens expansion*. Stockholm: SNS Förlag.

Margo, Robert A. 2004. Review of Peter H. Lindert Growing Public: Social Spending and Economic Growth Since the Eighteenth Century, Volume I: The Story. *Economic History Services, Apr 28, 2004*.

Ministry of Finance, Sweden. 2003. Government Proposition 2002/03:100, Bilaga 3.

Nelander, Sven and Ingela Goding. 2005. Anställningsformer och arbetstider 2005. *Ett faktamaterial om välfärdsutvecklingen Nummer 64* Löne- och välfärdsenheten, LO.

OECD. 2004. Benefits and Wages. *OECD Indicators 2004*.

_____. 2005. "Society at a Glance." *OECD Social Indicators - 2005 Edition.*

Persson, Torsten and Guido Tabellini. 1994. Is inequality harmful for growth? *American Economic Review* 84(3): 600-621.

Rones, Philip L. , Randy E. Ilg, and Jennifer M. Gardner. 1997. Trends in hours of work since the mid-1970s. *Monthly Labor Review* 120(4): 3-14.

Rothstein, Bo and Anders Lindbom. 2004. The Mysterious Survival of the Scandinavian Welfare States. *Paper presented at the annual meeting of the American Political Science Association, Chicago.*

Socialdepartementet. 2000. Socialbidrag 1999. *Socialtjänst 2000:8*.

Söderström, Lars, Anders Björklund, Per Gunnar Edebalk, and Agneta Kruse. 1999. *Från dagis till servicehus: Välfärdspolitik i livets olika skeden*. Stockholm: SNS Förlag.

ABOUT THE AUTHOR

Andreas Bergh received his PhD from Lund University in 2003, was visiting scholar at the Center for Basic Research in the Social Sciences at Harvard University in 2004, and is currently doing research at the Ratio-institute in Stockholm and at Department of Economics at Lund University. His research concerns the welfare state, its development, justification as well as issues regarding its consequences and how to measure them. Many of these issues are dealt with in his most recent publication "On the Counterfactual Problem of Welfare State Research: How Can We Measure Redistribution?" that appeared in *European Sociological Review* in 2005. A related research interest is human fairness behavior and the evolution of institutions such as property rights and social sharing norms. Email: andreas.bergh@ratio.se.

Go to Reply by Peter Lindert (2006)

Econ Journal Watch,
Volume 3, Number 2,
May 2006, pp 236-250.

The Welfare State Is the Wrong Target: A Reply to Bergh

PETER H. LINDERT *

I AM DELIGHTED TO HAVE THIS CHANCE TO EXPAND ON THE links between social spending and economic growth. Andreas Bergh (2006) raises both small and large concerns that deserve further exploration. Let us go first to the largest issues of scope and methods, then to his specific criticisms of parts of my *Growing Public* book (hereafter *GP*) relating to Sweden's policies toward the poor, toward women's work, and toward retirement. The final section of this reply invites him and others to re-focus their search for flaws in large government, since the welfare state, as actually practiced, has not become a major flaw. On the contrary, the social transfers that have always defined the welfare state are indeed a "free lunch" in the sense that they have delivered more equality and longer life expectancy at an essentially zero cost in terms of GDP. Rather, it is other forms of legal and governmental interference with markets that are more likely to be anti-growth.

Two Centuries, Many Countries

We now have the benefit of a rich international history for studying the sources and effects of social programs. My book has mapped much of that history, starting with a few European countries in the late eighteenth century and widening to a couple dozen countries from the late nineteenth century on. It covers all the historical experience in which democratic

* University of California, Davis. Davis, CA 95616.

countries devoted a large share of GDP to social spending, exploring why that social expansion came so late in history, and why it has come more to some countries than to others. The book also tackles the whole issue of the effects of tax-based social spending on economic growth. Bergh focuses on this side of the causal circle and on what he considers my "special use" of Sweden.[1]

TESTS VERSUS THEORIES

> *Think of a bumblebee. With its overly heavy body and little wings, supposedly it should not be able to fly—but it does.... This is how so-called analysts view the Swedish economy. We 'defy gravity.' We have high taxes and a large public sector, and yet, Sweden reaches new heights. We are still flying, so well that many envy us for it today.*
>
> —Prime Minister Göran Persson, Opening Address on March 10, 2000 to the Extra Party Congress of the Social Democrat Party in Stockholm[2]

Theory and fact are complements in the production of knowledge. Like all complements, they need to be mixed in the right proportions. Excessive fact-gathering calls for fresh theorizing. Excessive theorizing calls for fresh facts.

On the crucial issue of the welfare state and economic growth, theory has gone into overdrive. For some years now, there has been a widening gap between the empirical record and a story that is re-told with increasing insistence. No longer do we just hear that there is *the danger* that tax-based safety nets and anti-poverty programs *might* bring high economic costs. Notice how often we are told that economists have "shown" and "found" that this is true.

Such assertions are often a bluff. So theory-dependent are the latest parables about the ruinous welfare state that those words—"shown" and "found"—have been abused. Rather than showing or finding this result,

[1] *GP* devotes as much detail to Britain, or to the United States, or to the contrasts among 21 core OECD countries as it does to Sweden.
[2] As quoted in Thakur et al (2003).

they have *chosen to imagine* it.[3] The vulnerability of such theory-dependent bluffs to challenge has reached the point where a pro-state politician could rightly mock them before a partisan audience, as Prime Minister Persson did in the speech quoted above. My reading of the statistical and historical record leads me to very different tentative conclusions about the performance and prospects for the welfare state.

Andreas Bergh's critique is fortunately less theory-dependent than much of the literature pushing the high costs of taxes and social transfers. He has not been guilty of presenting simulations or pure models as fact. His arguments lean mainly on a detailed reading of Sweden's tax and benefits rules. Yet even the higher standard of evidence resting on the calculation of incentive wedges from a reading of statutes runs into a limit—a limit that has become obvious after thirty years of literature in the "I-see-a-tax-wedge" tradition. It is one thing to see a tax wedge, and quite another to show that it—and the larger package policy changes that accompany it—will bring a big effect in the real world. I return to this point below.

To compare theoretical nightmares with actual practice requires statistics and history. Given the large number of countries for which we now have the benefit of decades of fairly good data, there is no excuse for not working at length with multivariate testing of the competing hypotheses. Fortunately, many scholars have ventured the larger statistical voyage to find out how growth is affected by different structures of government taxes and expenditures. Most of the international econometric literature antedates my book. While my own work has presented what I consider more appropriate tests, my findings do not depart from the findings typical of the best-specified tests published by other economists.[4]

---

[3] This passage's contrast between what is "shown," "found," and "imagined" is elaborated in *GP* (Volume 2, Chapter 18, especially 82-84) and in my NBER working paper cited by Bergh.

[4] For the depth of the literature by other economists, see the surveys of it in Atkinson (1999) and *GP* (Chapter 18 in Volume 2, especially Table 18.1). Bergh hides the extent and thrust of this literature with some spin-control phraseology that seems designed to make it look as though my results are weak or un-representative: "Lindert has not been able to find a robust and statistically negative effect." "Lindert acknowledges that other studies find a negative effect...." (211). The only other econometric studies he cites on this negative effect are the Fölster-Henrekson study, which has been questioned in the other studies he cites, and the Gordon-Wang (2004) comment here in *Econ Journal Watch*. Neither of these studies is really about the welfare state or social transfers. The Gordon-Wang comment deals with a global cross-section, not a panel of OECD democracies, and thus mixes in behavior from such kleptocracies as Zambia and Zimbabwe. It finds an insignificantly positive effect of "scope of public sector" on economic development.

The current state of their econometric evidence is more interesting and more revealing than a reader might catch from Bergh's attempt to set it aside by saying, "It has long been known that it is hard to produce robust results regarding the effect of government size on growth" (212). A more accurate wording would be that most serious studies fail to confirm high net costs or high net benefits of big government among OECD democracies. That fact delivers some important information. 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. When repeated tests keep giving non-negative results about social transfers, one should be prepared for the possibility that the tests are telling us something about the real world. One should not keep repeating past assertions about the ruinous cost of the welfare state when the data refuse to reveal it—just as one should not just cite theoretical models saying that the whole package of social transfers is good for economic growth, which is also not supported by any clear test.

The path to firm conclusions requires both solid econometrics and some thinking about what real-world forces might explain the econometric results. The econometrics must meet several criteria. The samples must actually contain different policy regimes. The samples should also pertain to OECD democracies, and not to some global sample mixing too many structures together. The authors must grapple with the thorny issues of simultaneity, heteroskedasticity, and the non-linearity of GDP costs predicted by theory. The measure of government should actually refer to the social transfers of the welfare state. Tests that meet these criteria have contradicted the assertion that the overall welfare-state package has clear high costs.

A more persuasive study in this genre is Kneller, Bleaney and Gemmell (1999), which classifies taxes by whether they discourage investments, and divides expenditures into uses that are productive (general public services, defense, education, health, housing, transport, communications) or unproductive (social benefits, recreation, "economic services"). Using macro-panels with attention to budget accounting constraints, the authors reach intuitively persuasive conclusions: distortionary taxes (on income, payrolls, property) hurt growth if they finance unproductive expenditures, while non-distortionary taxes (on buying goods and services) do not, and productive government expenditures enhance growth, while unproductive ones do not. In Allard and Lindert (2006), we note that the Kneller-Bleaney-Gemmell study does not explicitly test for the welfare state and the implicit bundle of taxes that goes with it, a bundle tilted away from taxing capital and corporations, especially in Sweden's case. Their results seem consistent with those in *GP* (Chapter 18) that explore the relationship of transfers to tax mix.

Taking the next step of reflecting on how the statistics could have given such nullish results is particularly rewarding. One wades into the differences among different kinds of social transfers, different kinds of taxes, and differences in historical context. It turns out that there are good reasons why the net effect could be near zero, as most observers have concluded. Some social transfers are better than others, and some taxes are better than others. The "free lunch puzzle" remains as I described it when exploring these issues in *GP*, though much more remains to be done.

Tax Wedges versus Work versus Output

It is a long road from calculating an apparent tax wedge to delivering a final empirical result in terms of lost work or lost GDP. Perhaps one of the main reasons why Martin Feldstein and others who have contributed so well to the incentive wedge literature have not made the full trip to empirical results is that quantifying the wedges itself is an arduous task. Tax codes and social program rules are very complex. Their complexity raises the importance of doing incentive calculations, because society might be missing something by not knowing the possibly dangerous implications of how the various tax rates, exemption clauses, and economic shocks might interact. Yet the same complexity means that at the end of the exercise, one might have missed some important features of how a whole policy regime change might play out. One common misstep is to fail to look at the full budgetary implications of a particular fiscal shift. Studies of the effects of marginal taxation talk of their "burdens" without exploring how productively taxes are spent. Studies of the distribution of social "benefits" often gloss over the effects of the taxes that pay for them. The danger of missing something is especially great if one *chooses* to look at only part of the picture and then presents it as the whole picture.

For calculated tax wedges, one must sooner or later look at actual performance of the entire national economy to determine how it affects work and output. Let us turn first to the work effect. Bergh, like anybody else worried about what look like high statutory marginal tax rates, must struggle with a severely limiting fact: There is no negative international relationship between employment rates and the main welfare-state indicators, such as the share of transfers in GDP or the statutory generosity

of unemployment compensation.[5] Nor is the employment ratio lower in the countries for which we have reasonable estimates of overall marginal tax rates. Bergh acknowledges the same point in his two-country comparison of employment ratios in Sweden and the United States.

There are two basic ways to reconcile high statutory tax rates with the absence of work and output response. One is to conclude that people have such low elasticities that the marginal tax rates simply don't matter. The other possible resolution is that the high marginal tax rates are misleading in the first place. The second resolution seems more plausible. Many high statutory tax rates are seldom paid, often because other features of the tax system were designed to allow people to avoid those high rates.[6] In addition, people see offsetting benefits of work within the welfare state. That the benefits of work within the high-budget welfare state are missed by those analysts intent on seeing taxes as merely taxes is well illustrated by the OECD measures of tax wedges on employment. A large share of these consists of mandatory contributions to the social security system. Yet the individual's benefits from that same system are tied to those paycheck deductions, even if the benefits are "pay as you go" in terms of entire cohorts.

Even if a given tax incentive cuts work, the output results might well be smaller than the work effect. If the higher marginal tax cuts labor supply, employers may well respond by substituting capital and shifting toward the types of labor less affected by the higher wedge. That is, the same mainstream theories that predict a work loss should predict a smaller output loss in percentage terms. The empirical elusiveness of output losses from the welfare state may relate to such productivity offsets, though the prevailing empirical tendency has already been noted: The employment effects themselves are hard to distinguish from zero.[7]

---

[5] Using data for 21 OECD countries from 2001, for example, the correlation is zero (0.021) between the overall employment ratio for persons 15-64 and social transfers as a share of GDP. The correlation is slightly negative for men and slightly positive for women.

[6] The built-in offsets to what look like high marginal rates occur both for high- and low-income households in Sweden. I return later to the ways in which the rich never faced an average effective marginal tax rate as high as the top rates before the 1991 reform. Even for low-income recipients of welfare transfers, the incentives were often less than meets the eye, as the Lindbeck Report, which wanted to stress that Sweden's tax burdens are excessive, briefly cautioned its readers: "[t]he size of public-sector spending [and taxation] in Sweden is not strictly comparable to that of many other countries, because several types of transfers in Sweden are taxed, while in many other countries they are not" (Lindbeck et al 1994, 5).

[7] An exception is the econometric literature that uses the OECD's *employment* tax wedge as its tax variable. That literature does find a negative effect on employment, but does not follow

Tax Wedges at the Bottom and the Top

Bergh announces very high marginal rates of effective taxation at the bottom and top of the Swedish income ranks, contradicting my argument that the rates have been smoother than in Britain or America. This is an important issue well raised. Let us turn first to the single mothers at the bottom of the ranks, and then to the tax rates faced by the richest.

Bergh's case of statutory tax rates and benefit loss rules facing single mothers is well chosen, since he is talking about a passage in my book (Table 10.2 in Chapter 10 of Volume 1) where I too talk about the marginal rates, rather than about the behavioral response. He 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. I had used other studies that implied lower rates, in particular the Gustafsson-Klevmarken (1993, 106-110) calculations of a 30 percent marginal rate for a single parent and for someone liable for child support payments. Other studies get varying results, and some do indeed calculate high marginal tax rates like those that worry Bergh. For example, the recent IMF study of Sweden worries about the same issue, though it does not take as firm a stand as Bergh does (Thakur *et al.* 2003, esp. 48-50). His claiming a lower marginal rate for a single mother in the United States than I had presented, based on the work of Acs and others, seems to misunderstand which era in American history I was discussing. He is talking about the United States since the generosity of the Earned Income Tax Credit (EITC) was dramatically raised in 1993, whereas I had contrasted Sweden mainly with the United States policies of the 1980s. For that earlier era, before the jump in EITC, the universality contrast between countries still looks the same, adjusted for EITC or WFTC, especially if the provision of health care, child care, and other services is less means-tested in Sweden than in America.

Higher up the income spectrum, Bergh seems to contradict a considerable literature without any solid evidence.[8] He asserts that "the

through and explore whether such taxes specific to jobs have a negative effect on output. Unreported regressions in the project by Allard and Lindert (2006) confirm that specifically taxing employment reduces employment, but does not significantly reduce output.

[8] He cites only his own 2004 article. That article does not present any fresh evidence on top marginal income tax rates, nor does it rebut the earlier evidence I cite in the next footnote. Rather, his 2004 article is mainly just a sensible and readable discussion of the different meanings of universality, which he concludes did not decrease in the crisis of the early 1990s.

structure of the marginal net tax rates in Sweden is exactly the opposite of Lindert's description—high for low and high-income earners, and lower for middle income earners" (219). While there lingers some uncertainty about those low-income single mothers, he surely owes us an explanation of why he passed up all the evidence that top income earners did not pay much higher rates than middle earners, either before or after the tax reform of 1991.[9] 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. I had pointed all this out in *GP* (Chapters 10, 11) in the hope of encouraging a new look at how tax structure relates to the welfare state. And if the marginal tax and benefit-loss rate were really so high, Bergh would again need to present evidence that it actually affects labor supply more than in other countries.[10]

Women's Hours

Having acknowledged that Swedish women have a higher relative pay than in other countries and a higher employment ratio than in most other countries, including the United States, Bergh goes through some contortions to dismiss that achievement and my explanation of it. My explanation was that so many of them had jobs because Sweden paid women relatively better than in other countries, and women respond with a relatively high elasticity of labor supply.

Bergh disagrees, on the grounds that the average employed woman worked shorter hours in Sweden than in the United States, so that "[w]hen actual working hours are taken into account, we find that Lindert is mistaken about Sweden having higher female employment" and "they spend more time in unpaid non-market work" (220). He thus implies that women's formally working 10 percent fewer hours on the average would

[9] Hansson and Stuart (1990); Gustafsson and Klevmarken (1993, 64); Aronsson and Walker (1997, 236-238); Norrman and McLure 1997; and Thakur *et al.* (2003), showing that the marginal effective rate of labor taxation did not peak at the top and that Sweden's marginal effective rate of capital taxation is low relative to other countries.

[10] Denmark's case in the 1980s and 1990s was even more extreme than Sweden's. For a single mother at the bottom of the income structure, the statutory marginal tax rates looked even higher than those posted by Bergh. Yet the in-depth studies found almost no labor supply response at all, either for this sub-group or for the labor force as a whole (Pedersen 1993, especially 273-288; Mogensen 1995).

cancel the work benefits of having jobs for 10 percent more women. His overall context invites readers to think that Swedish women turn to part-time work because they face higher marginal taxes—despite the fact that from 1963 to 1992, as the statutory marginal tax rates probably climbed, women's average hours per week did not decline as men's hours did.[11]

His imagery here is implausible. By equating a percentage difference in hours with a percentage difference in productivity, he implies that the value of an hour of a woman's extra home time is worth zero. 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? Unlike their American counterparts, they are not under pressure to keep a full-time job just to keep their health insurance.[12]

Retirement

For the age of retirement, we confront the same clash between calculated marginal tax rates and employment results as for the wedges discussed earlier. For men who would reach the ages 55-64 in 1985-1994, the Gruber-Wise calculations do imply that Sweden gives senior men a stronger incentive to retire than in the United States (Gruber and Wise 1999, *passim*).

Yet for retirement as for all other labor margins, we must compare the international differences in a synthetic tax-rate calculation with the actual work outcome. Among the 21 core OECD countries in 1999, a greater share of men 55-64 were still at work than in most other countries. Sweden ranked fifth in this respect, and the United States ranked sixth. For women, the share still working in the 55-64 age bracket is also high by

[11] Aronsson and Walker (1997, 259). For more up-to-date hours estimates see either OECD or the Groningen estimates online.

[12] Bergh offers no reason for saying that he is "inclined to reject Lindert's elastic labor supply explanation of Sweden's high female employment ratio" (221). He has not denied that the relative pay favors women more in Sweden than elsewhere. As for the elasticities of labor supply, the literature is clear that it is higher for women than for men, other things equal, in Sweden and elsewhere. See, for example, Aronsson and Walker (1997, 239-244) and Gustafsson and Klevmarken (1993, 77).

international standards, despite Bergh's claims to the contrary. Again we face that choice between two possible reconciliations: Does Sweden's favorable work effort among seniors belie the high marginal tax rates because the elasticity is inherently low, or because the actual net incentive is not as negative as the calculated tax rates suggest? Again, the latter reconciliation is close to my hunch, but either way the issue must be explored, and just calculating an apparent statutory tax wedge is not sufficient to explain actual behavior.

Public versus Private Jobs

In a revealing quip, Bergh tips off his readers that "Lindert. . . . does not tell us that in Sweden, there has been no net growth of jobs in the private sector since 1950," reproducing a graph from Davidsson and Henrekson. Bergh's point about public sector jobs stops there, as if to say "Q.E.D." (221-222).

Yes, the rise in jobs is virtually all in the local-government sector, and virtually all female. Dominant are health-care professionals, educators, and day care providers. But who is the reading audience here? Are readers supposed to shake their heads in knowing disapproval, because they all know that public employment is inferior to private? Again, we are back to that divide between those who are willing to be swayed by direct tests of effects on outputs and wellbeing and those who want to hear only what fits their pre-conceptions.

Changes versus Levels in "Freedom"

Bergh talks against himself in stressing that Sweden's system was reformed in the 1980s and 1990s. 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: "[T]he reforms toward increased economic freedom seem to be a good candidate for explaining the survival of the Swedish welfare state" (233). He lists a host of reforms in separate sectors of economic life. This is somehow supposed to contradict Lindert, who "does not attempt to describe" the wave of reforms and its connection with the survival of the welfare state. Yet I did in fact discuss the reforms, and even more reforms than his final page credits me with. I also discussed some reforms in other countries, particularly the major reform of disability

claims in the Netherlands.[13] 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.

The Welfare State Is the Wrong Target

Bergh's comment illustrates a puzzling tendency in the debate over big government and the welfare state among OECD democracies. For some reason many critics still prefer to attack the welfare state as if it were some all-purpose Leviathan dragging down freedom and growth. Yet it has not dragged down economic growth, and of all the forms of government intervention social transfers are not the ones that infringe on the flexibility of free markets the most.

To underline the importance of changing targets, let me share Bergh's chosen emphasis in the trans-Atlantic contrasts in growth factors. That is, let me concentrate here on European institutions that have probably checked the advance of growth and well-being, rather than such American inefficient-policy suspects as its health care/insurance system, its under-investments in the human capital of the poor, and a tax structure that puts too much emphasis on taxing capital and not enough on taxing gasoline, alcohol, and tobacco.

Other European institutions deserve the preliminary indictments that have been misdirected at the welfare state. The same international-panel econometric method that fails to indict the welfare state does indict other institutions. One culprit in particular has been Europe's employee protection laws (EPLs) dating from the late 1960s and 1970s. By blocking firing, these have also blocked hiring and have probably slowed the rate of re-allocating labor to rising sectors. What is especially clear about EPLs is that they have created insiders and outsiders, much as the model by Lindbeck and Snower (1988, 2001) describes. The negative effect of EPLs is likely to show up only with a lag, and in the form of eventually reduced productivity, since it will take some years for the human capital loss from outsider status to outweigh the human capital gains for the insiders. For as long as the econometric literature was testing for the effect of EPLs on the overall unemployment rate, its results were not fully robust. Yet testing for

[13] *GP*, pages 271-282, 288-291, 298-300. As for Bergh's implying that I overlooked school vouchers, I did discuss them and the evidence on their merits in four other European countries and Chile (*GP*, 165-167), though I did not mention Sweden.

the eventual productivity losses seems to have strengthened the case again EPLs.[14]

The locus of the problem of EPLs within Europe differs from the locus of the welfare state. EPLs seem to delay jobs and growth mainly in Mediterranean Europe, whereas Italy and Greece are not particularly generous with welfare or unemployment compensation or active labor market policies. While Sweden and the other Nordics have EPLs, they are not as strict as in the Mediterranean, and this seems to have shown up in productivity as well as in the relative unemployment of youths and women.

A second European institutional drawback is in the degree of product market regulation. Here the OECD measures for seven large sectors seem to tell a plausible story, even though they miss the likely importance of restrictions on retailing. Product-market liberalization has an interesting geography. The lead was taken by the Pacific Rim countries (especially New Zealand), with Mediterranean Europe again being the slowest to liberalize. Econometric tests seem to indict product market regulations as a further source of productivity loss in Europe.[15]

Finally, a particular American advantage over Europe (and Japan and Latin America) lies in policy toward higher education. Among the social sectors, higher education is perhaps the one where ordinary competition works best. While there is a case for subsidies to correct for knowledge externalities, there is no case for centralized work rules or for dominance of government funding. It is specifically the Americans who have made the least error in this social sector. The public share of funding is lower in American higher education, and state universities must compete with each other as well as with private universities, for research funding, for faculty talent, and for student talent. Students' consumer evaluations also play as strong a role on U.S. campuses as in any other country. Sweden's system seems to be in better shape than those of countries to the south (e.g. Italy), but it is not clear what gains Sweden gets from having centralized university education, and in having made it tuition-free in the past.

Overcoming the excessive targeting of the welfare state sets the stage for a larger inquiry plotting the efficiency boundaries between markets and governments.

---

[14] Allard and Lindert (2006).

[15] Allard and Lindert (2006, Table 2). While our product-regulation result seems plausible, I should add that it is not necessarily robust to one's choice among defensible regression specifications.

REFERENCES

Allard, Gayle J. and Peter H. Lindert. 2006. Euro-Productivity and Euro-Jobs: Which Institutions Really Matter? UC-Davis, Department of Economics Working Paper no. 06-19. Link.

Aronsson, Thomas and James R. Walker. 1997. The Effects of Sweden's Welfare State on Labor Supply Incentives. In *The Welfare State in Transition: Reforming the Swedish Model*, ed. Richard B. Freeman, Robert Topel, and Birgitta Swedenborg. Chicago: University of Chicago Press for the NBER, 203-265.

Atkinson, A.B. 1999. *The Economic Consequences of Rolling Back the Welfare State*. Cambridge: MIT Press.

Atkinson, A.B. and Gunnar Viby Mogensen, eds. 1993. *Welfare and Work Incentives: A North European Perspective*. New York: Oxford University Press.

Bergh, Andreas. 2004. The Universal Welfare State: Theory and the Case of Sweden. *Political Studies* 52: 745-766.

_____. 2006. Is the Swedish Welfare State a Free Lunch? *Econ Journal Watch* 3(2): 210-235. Link.

Gordon, Peter and Lanlan Wang. 2004. Does Economic Performance Correlate with Big Government? *Econ Journal Watch* 1(2): 192-221. Link.

Gruber, Jonathan and David Wise, eds. 1999. *Social Security Programs and Retirement around the World*. Chicago: University of Chicago Press.

Gustafsson, Björn and N. Anders Klevmarken. 1993. Taxes and Transfers in Sweden: Incentive Effects on Labour Supply. In *Welfare and Work Incentives: A North European Perspective*, ed. A.B. Atkinson and Gunnar Viby Mogensen. New York: Oxford University Press, 50-134.

Hansson, Ingemar and Charles Stuart. 1990. Sweden: Tax Reform in a High-Tax Environment. In *World Tax Reform: Case Studies of Developed and Developing Countries*, ed. Michael J. Boskin and Charles E. McLure, Jr. San Francisco: ICS Press, 127-148.

Kneller, Richard, Michael Bleaney, and Norman Gemmell. 1999. Fiscal Policy and Growth: Evidence from OECD Countries. *Journal of Public Economics* 74(2): 171-190.

Lindbeck, Assar. 1997. The Swedish Experiment. *Journal of Economic Literature* 35(3): 1273-1319. Expanded version: *The Swedish Experiment*. Stockholm: SNS Förlag, 1997.

Lindbeck, Assar, et al. 1994. *Turning Sweden Around*. Cambridge: MIT Press.

Lindbeck, Assar and Dennis J. Snower. 1988. *The Insider-Outsider Theory of Employment and Unemployment*. Cambridge: MIT Press.

_____. 2001. Insiders versus Outsiders. *Journal of Economic Perspectives* 15(1): 165-188.

Lindert, Peter H. 2004. *Growing Public: Social Spending and Economic Growth since the Eighteenth Century*. Two volumes. Cambridge: Cambridge University Press.

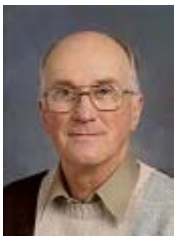
Mogensen, Gunnar Viby, ed. 1995. *Work Incentives in the Danish Welfare State: New Empirical Evidence*. Aarhus: Aarhus University Press.

Norrman, Erik and Charles McLure Jr. 1999. Tax Policy in Sweden. In *The Welfare State in Transition: Reforming the Swedish Model*, ed. Richard B. Freeman, Robert Topel, and Birgitta Swedenborg. Chicago: University of Chicago Press, 109-153.

Pedersen, Peder J. 1993. The Welfare State and Taxation in Denmark. In *Welfare and Work Incentives: A North European Perspective*, ed. A.B. Atkinson and Gunnar Viby Mogensen. New York: Oxford University Press, 241-288.

Thakur, Subhash, Michael Keen, Balázs Horváth, and Valerie Cerra. 2003. *Sweden's Welfare State: Can the Bumblebee Keep Flying?* Washington: International Monetary Fund.

ABOUT THE AUTHOR



Peter H. Lindert is Distinguished Professor of Economics at the University of California, Davis. His books and journal articles have dealt with modern inequality trends, the welfare state, human fertility, international debt crisis, international trade competition, land quality, farm policy, soil history, and other topics. His latest book, *Growing Public: Social Spending and Economic Growth since the Eighteenth Century* (two volumes, 2004) has been covered in the *Economist, Newsweek*, the *New York Times* and other media, and has received both the Allan Sharlin Award for Best Book in Social Science History and the Gyorgy Ranki Prize for best book in European Economic History. He has served as the elected President of the Economic History Association. His email address is phlindert@ucdavis.edu.

Go to Comment by Andreas Bergh (2006)

Econ Journal Watch,
Volume 3, Number 2,
May 2006, pp 251-274.

Farley Grubb's Noisy Evasions on Colonial Money: A Rejoinder

RONALD W. MICHENER AND ROBERT E. WRIGHT *

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

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

GRUBB'S RECENT PAPERS (2003, 2004, 2006B) ARE AIMED AT nothing less than rewriting important chapters of early American history. Our goal in both our *AER* and *EJW* comments was a negative one, to dissuade readers from accepting Grubb's views and data. We are humbled by the complexity of the early American monetary system and the meagerness of the available evidence. We nowhere make the blanket claims that Grubb attributes to us, namely, that specie was plentiful, that exchange rates were immutable, or that cross-colony circulation of bills of credit was ubiquitous. Instead, we offer evidence that at some times and places specie was more abundant than Grubb claims, that most colonial exchange rates oscillated within broad specie points, and that bills of credit often circulated in adjacent colonies. And again, we make such claims with one point only in mind, to alert scholars that Grubb's interpretation is highly suspect.

* Michener: Department of Economics, University of Virginia. Charlottesville, VA 22904. Wright: Department of Economics, Stern School of Business, NYU. New York, NY 10012.

Our central disagreement is very simple. We believe that when colonists promised "six shillings Pennsylvania money" they usually meant "I will give you bills of credit (not necessarily Pennsylvania's), gold, silver, tobacco, hogs, credit in my account book, etc. to the value of six shillings." Grubb believes that in certain transactions "six shillings Pennsylvania money" invariably meant "I will give you six shillings in Pennsylvania bills of credit." Grubb's supposition that pounds, shillings and pence in runaway ads must refer to Pennsylvania's bills of credit, has, in our view, been thoroughly undermined. Grubb never explains why there were profuse references to Pennsylvania pounds and shillings in runaway ads before there were Pennsylvania bills of credit; his argument that ads promising specie were legally binding commitments to pay in specie is convincingly refuted by Pennsylvania's legal tender laws; he never explains why his technique, when applied to the early Federal period, produces results that contradict his own work in Grubb (2003, 2006b). A host of other evidence could be presented—advertisements and contemporary documents that simply make no sense if pounds and shillings meant only paper money. The following call for a subscription to a book is one such example:

> II. The Price to Subscribers will be twenty Shillings each Book, stitched in Marble Paper. III. One Dollar Advance, being the first Part of the Subscription Money, to be paid at the Time of subscribing; the other twelve Shillings and Sixpence to be paid at the Delivery of the Book. (*Pennsylvania Evening Post*, 11 May 1775)

Such evidence he labels "anecdotal" and ignores. (Or, he contrives unlikely scenarios to explain them, like the bookseller sought to engage in currency speculation.) Our view effortlessly explains such seeming anomalies, which are more common than Grubb concedes.[1] "Dollar" and 7 shillings 6 pence were synonymous in late colonial Pennsylvania because they referred to the same amount of economic value, not to specific physical exchange media. We did not concoct this "doomsday weapon," but merely applied McCusker's dictum (1978, 3-6) that one should not conflate units of account with media of exchange.

In the lengthy appendix here, we address the remainder of Grubb's reply. We relegate the details to an appendix because we believe that, for a

[1] For additional examples, see Michener and Wright (2006b, 25).

general audience, the tone of Grubb's reply is the best evidence of the underlying weakness of his propositions.

Appendix:

REMARKS PERTAINING TO OUR REASONS FOR DISBELIEVING GRUBB'S TIME SERIES

Our legal tender evidence

In our comment, we note that Grubb (2004, 333) describes runaway ads as legally binding commitments to pay in the media advertised. That, he argues, is what guarantees that the ads reflect the medium of exchange tendered. We (2006a, 9-10) reply that all Pennsylvania bills of credit issued before 1764 were a full legal tender even in contracts that specified payment in specie. Ads, therefore, could not have been legally binding commitments to pay in the advertised media. Grubb makes no comment on the statutes or the Privy Council deliberations they provoked, although we (2006a, 9-10 fn. 3, 23-24 fn. 15) discuss both. However, in direct contradiction to our evidence, he (2006a, 57) reiterates that "most quit rents . . . were required to be paid in sterling (specie)."

On the similarity of ads before and after 1723

We (2006a, 7) present runaway ads from the *American Weekly Mercury*, printed in 1720, three years before Pennsylvania issued its first paper money. Those ads, like the overwhelming majority printed in the *Weekly Mercury* that year, offer rewards in undesignated pounds and shillings. What reason is there to believe, we ask, that pounds and shillings in similar ads printed after 1723 must invariably mean Pennsylvania's paper money? Grubb did not respond.

On the robustness of Grubb's procedure

We (2006a, 10-11) note that Grubb has used two kinds of arm's-length transactions to determine the ratio of specie to paper money transactions: redemptioners' contracts (used in Grubb (2003, 2006b) to infer the composition of Pennsylvania's money supply in the confederation and early federal eras) and runaway ads (used in Grubb (2004) to infer the composition of Pennsylvania's colonial money supply). We point out that if you reverse the procedure and use redemptioners' contracts to infer the composition of the colonial money supply, or runaway ads to infer the composition of the confederation and early federal money supply you obtain wholly inconsistent estimates.

Grubb (2006a, 68) responds to this point. The colonial records for redemptioners, he writes, were "recorded by the [Philadelphia] mayor's office as a contract registration exercise," while the confederation records were "recorded under the auspices of the German Society of Pennsylvania as an honesty-in-contracting monitoring device." Therefore, he continues, he "would not expect them to reflect monetary usage in the same way . . . ." He does not, however, explain why not.

Grubb's supposition, we thought, was that *in arm's-length transactions between strangers the unit of account accurately records the medium of exchange*. Grubb notes correctly that we ourselves do not have a model of the unit of account.[2] The unit of account, in our opinion, arises as some complex admixture of convenience and tradition; even today anomalies exist that would be difficult for any model to fully explain, such as the role guineas play as a unit of account in modern British horse racing. Solving this complex problem, however, is not a prerequisite to pointing out the shortcomings of the data series arising from Grubb's supposition. We believe we decisively refuted Grubb's supposition with our comparison of colonial redemptioners' records and runaway ads. Grubb replies that we have misconstrued his supposition—who records the transaction and why is also important. We believe this renders his supposition too vague and ill-defined to be of much use. Why should data derived from newspaper ads placed by thousands of anonymous Pennsylvanians reflect the medium of exchange if contract registrations at the Mayor's office do not? Nor is this

[2] Grubb (2006a, 52-53) writes: "What is missing in Michener and Wright's analysis is any model for determining what money the unit of account will be in, when and why this unit will shift to being a different unit of account, when and why multiple units of account will exist in society at large or within the same individual transaction, and so on."

the only conundrum. Runaway ads in the *Pennsylvania Gazette* during the early Federal period were overwhelmingly denominated in dollars at a time when the redemptioners' contracts were overwhelmingly recorded in pounds and shillings, as demonstrated in Table 1 below.

Table 1: Units of account in the early federal era[3]

Year	*Redemptioner's A Contracts* Percent pounds	*Pennsylvania Gazette* Runaway ads Percent pounds	Z-statistic testing Equality of proportions.
1790	93 (n=39)	12 (n=17)	5.9
1791	100 (n=151)	23 (n=13)	11.1
1792	99 (n=270)	33 (n=24)	12.5
1793	100 (n=153)	14 (n=21)	12.1
1794	100 (n=200)	13 (n=32)	14.1
1795	96 (n=367)	6 (n=17)	13.3

Source: Michener and Wright, 2006a, 11; 2005, online appendix

Yet in analyzing the confederation and early federal era, Grubb (2003, 2006b) bases his estimates on these redemptioners' contracts. If *Pennsylvania Gazette* runaway ads in the colonial period reliably reflect the medium of exchange, why aren't the ads placed during 1790-1795 equally reliable? Grubb did not respond to this point. What exactly is Grubb's supposition?

Parenthetically, a glimpse of the composition of the colonial money supply can be obtained by examining rewards for lost or stolen money because such ads often explicitly listed the precise exchange media for which a reward was offered. The emphasis is on the lost/stolen monetary items, not on the reward, and unlike runaway rewards, these ads often suggest the medium of exchange was remarkably diverse. Consider, for example, the following advertisement from the *Pennsylvania Packet* (6 September 1773):

[3] We used the online edition of the *Pennsylvania Gazette* and obtained our sample by searching for the string "ran away." Only advertisements placed by Pennsylvania residents were used. Duplication was avoided as far as possible. Advertisements mentioning multiple runaways were counted only once. Small rewards less than $1 (7 s. 6 d.) were not counted. Such advertisements (commonly 6 d.) were placed to insult the servant, not to induce his or her capture and return. Decimalized insults (commonly, $0.06) did not come into use until late in the period.

> Philadelphia, August 30, 1773. LOST, . . . on the road leading from Philadelphia to Bristol, and across the river Delaware to Burlington … A BUNDLE of paper currency and gold coin . . . the money is four Jersey bills of 3l. six ditto of 1l. 10s. five ditto of 15s. one York bill of 2l. and a Maryland two-thirds of a dollar; a heavy half Johannes, and two quarter ditto, one of them plugged, and one moidore . . . .

Transactions velocities

We (2006a, 11-12) point out that Grubb assumes that specie and paper money had the same transactions velocity in colonial times. Yet, in an earlier exchange, we pointed out (Michener and Wright, 2005, 686) that only $0.06 per capita in Pennsylvania's state-issued paper money survived into 1794—a year in which several dollars of specie and banknotes per capita circulated—yet Grubb asserts virtually all transactions were in state paper money. For paper money to have been a small fraction of the money supply, but to have executed nearly all transactions, it would have had to circulate much faster than specie and banknotes; we estimate, about 10,000 times faster. Why, we wondered, is it plausible for Grubb to argue that specie and paper money circulated with the same transactions velocity in colonial times, when he elsewhere contends that the transactions velocities were very dissimilar in the confederation era?

Grubb (2006a, 67) rebukes us for putting words into his mouth: "I did not claim that the velocity of circulation of Pennsylvania paper money in the late 1780s was higher than that of specie (Grubb [2005], 1343). I only said that people at the time seemed to think the velocity of circulation of Pennsylvania paper currency in the late 1780s was quite high."[4]

[4] True, but not germane. In 1786 Robert Morris remarked that while the quantity of state-issued paper money outstanding was small,

> what security have we that the next house of assembly will not issue another emission? The doubts and fears of this, and of tender laws, destroy the confidence of the public. While these doubts remain in the minds of the people, the circulation of paper must necessarily be quick, as no one will risque the keeping it long by him. (Carey, 1786, 42-43)

There are two difficulties with Grubb's argument, however. First is that this testimony to the public's lack of faith in state-issued bills is anathema to Grubb (2003), and that those who noted paper money's high velocity in the late 1780s were commonly (like Morris) its foes,

Grubb is correct that he never explicitly compares transactions velocities in the confederation era. However, in Grubb (2005, 1343, fn. 6) he cites Hepburn: "Between 1790 and 1795, the total volume of dollar money per capita in the United States increased 159 percent—from $3.00 to $7.77." Grubb goes on to attribute the rise in Pennsylvania prices between 1790 and 1795 to this increase; in short, Grubb is willing to accept Hepburn's estimate. Once one accepts Hepburn's estimate of the total money supply, or any estimate of this order of magnitude, the difference in transactions velocity we allude to is an inescapable implication.

We also noted that Pennsylvania's colonial bills of credit were disproportionately in small denominations compared to the gold and silver coins in circulation, and that in the modern era, small denomination currency circulates more rapidly than large denomination currency. It is plausible to think the same was true in colonial times. Grubb did not respond to this point.

Pennsylvania's Money Supply in the late 1740s and early 1750s

The issue here is very straightforward. Grubb's method suggests there was very little specie in circulation in Pennsylvania during this era—on the order of 4,317 Pennsylvania pounds. A number of contemporary documents suggest a much larger figure—on the order of 300,000 - 400,000 Pennsylvania pounds. We present our sources in Appendix 1 of our original comment, and Grubb (2006a, 57-60) berates their reliability. Grubb's lengthiest objections pertain to the fourth and final document, a 1753 letter (Hockley 1753) from Penn's quit rent collector reporting money was not so scarce as pretended, and citing as proof the fact that "full four fifths" of the money received in payment of quit rents was "Gold and Silver." We largely anticipated Grubb's objections and preemptively responded to them (2006a, 23, fn. 15). Since we feel that response is largely adequate, we shall

the very people that Grubb (2005, 1342) dismisses as "passionate, polemical and propagandistic." Second, by 1794 (the year in question) circumstances had changed. Ratification of the Constitution dispelled anxiety that more bills might be issued or that existing issues would be declared a tender. Pennsylvania's 1785 emission reached its nadir in 1788, having lost by that time one third of its value. However, beginning in 1788, the bills began to appreciate. By 1794, Pennsylvania had redeemed practically all of them at par, and no one seems to have doubted that the remainder would also be redeemed. See Michener and Wright (2005, 688-690).

focus on his other objections and relegate the details to a footnote.[5] Our

[5] Any reader hardy enough to attempt to get to the bottom of this exchange is going to find it difficult, because the give and take have not been reproduced in chronological order. The dispute over this passage began when we drew it to Grubb's attention in private correspondence. Grubb wrote a rebuttal which appeared in Grubb (2001), but not the published version. Grubb has seen fit, however, to reproduce that rebuttal without any significant alterations. It begins in the middle of page 57 in Grubb (2006a) and continues to the middle of page 58. The central point of that rebuttal is (Grubb 2006a, 57) that "most quit rents . . . were required to be paid in sterling (specie)." We responded in Michener and Wright (2006a, fn. 15). We noted that the Pennsylvania bills of credit were a legal tender even in contracts that specified payment in specie, and quoted from the statute book. We also noted that, much to the annoyance of the proprietor, the Pennsylvania courts were forcing him to accept bills of credit at their face value (i.e. as proclamation money) in the payment of sterling quit rents. This practice denied him about 20% of the value of his quit rents. The proprietor responded, in 1732, by forcing new tenants to sign contracts requiring them to pay either in specie or in paper money at its actual value. In short, the proprietor did not object to accepting paper money, only to accepting paper money at an artificially enhanced value. That contract, however, was unenforceable under the terms of the legal tender act. In 1739, the assembly responded by making a compensatory payment to the proprietor. In exchange, the proprietor agreed to accept paper money at its artificially enhanced value on grants made pre-1732, whereas those with post-1732 grants were to pay their quit rents "according to the tenor of the grants." Post-1732 tenants, like pre-1732 tenants, were unambiguously permitted to pay quit rents in bills of credit. All this is unmistakably clear from the primary sources quoted in Michener and Wright (2006a, fn. 15). What is not so clear is whether the Pennsylvania courts (after 1739) treated the 1739 agreement as an amendment to the legal tender act. That is, when a post-1739 tenant paid in provincial bills of credit, was he forced to pay the full sterling value of the quit rent, or did the legal tender act still permit the tenant to pay a lesser value? In 1760, the proprietor, dissatisfied that he had not received adequate compensation for his losses, appealed to the Privy Council and the Privy Council ordered that quit rents be excluded from future legal tender clauses. We were unable to determine to our satisfaction, however, whether the proprietor was complaining about all grants or only pre-1732 grants. We pointed out that one historian, Hutson (1970, 431), believes the answer is all grants. The controversial portion of Hutson's remarks (which describe Pennsylvania assembly deliberations in 1764) are reproduced below:

> For decades the Penns had made their tenants sign contracts pledging to pay their quit rents in either sterling or provincial paper money at the rate of exchange between Philadelphia and London and for just as long the Assembly had thwarted them by issuing legal tender paper currency with which the inhabitants discharged their obligations. By authorizing Penn to refuse to accept legal tender paper, the Privy Council had resolved [in 1760] the long conflict in his favor.

This leads to Grubb's final round of objections (Grubb 2006a, 58). Hutson's statement, Grubb asserts, could refer only to pre-1732 contracts. We disagree. Hutson says that tenants paid their quit rents in paper under the sanction of the legal tender act for decades *even though the proprietor had insisted they sign contracts to pay in either sterling or provincial*

first document was an anonymous Massachusetts pamphlet, written in 1749, that very explicitly states that the medium of exchange in New York and Philadelphia was silver and that the proportion of paper bills was small compared with silver. Grubb objects that the pamphlet was written as a polemic against paper money and that the author may have had no first hand knowledge of conditions in New York and Pennsylvania, and that it therefore "lacks credibility." We disagree, but there is little to say except that the pamphlet gives the impression of being temperate and well-written.

As for our other sources, Grubb (2006a, 59) writes: "Finally, Michener and Wright (2006a, Appendix 1) present evidence that the governor of Pennsylvania in the early 1750s resisted approval of new paper money emissions by the Pennsylvania Assembly because he thought there was lots of specie in the colony and so paper money was not needed. Michener and Wright take this evidence at face value." Grubb then relates, accurately, how the proprietor actually wanted more control over spending the paper money authorized, but that since the governor viewed it as impolitic to reveal the proprietor's true motives, he objected to the bill only on the grounds that specie was sufficiently abundant so as to make more

paper money at its actual value. The proprietor only began writing such contracts in 1732, so the only payments that could have violated such contracts were those written post-1732.

Grubb notes that if the inhabitants discharged their post-1732 quit rents with paper money, as Hutson maintains, Hockley is incorrect to say that four-fifths of all payments were received in specie. Moreover, Grubb points out, only an irrational fool would overpay his or her quit rents by tendering specie if it was financially disadvantageous to do so. We recognized the problematic nature of Hutson's analysis when we cited it, which is why we wrote: "*If Hutson (1970) is correct* . . . [about post-1732 contracts]" One need not impugn Hockley or declare the residents of Pennsylvania fools to reconcile these matters; there are at least two sensible explanations that preserve Hockley's integrity and the rationality of Pennsylvanians. Hutson may be wrong about post-1732 contracts, and the specie paid to Hockley could have arisen from payments on post-1732 contracts. Even if Hutson is correct about post-1732 contracts, Hockley might well have been a sensible individual. Knowing that the legal tender provisions forced him to accept roughly $0.80 on each dollar of quit rents owed, it would be sensible for him to accept an equivalent value in specie in lieu of paper. Our hunch is that this is precisely what Hockley did. Grubb doesn't explain what motive Hockley would have to lie about this matter in a private letter to the Proprietor, and Hockley clearly used this observation to illustrate how much more plentiful specie was than paper money.

Finally, accepting our position on the legal tender laws does not imply that property sellers detailing sterling obligations in newspaper advertisements were "knowingly writing nonsense." Even the tender laws required one to pay four Pennsylvania pounds for every 3 pounds sterling owed, which is ample reason for sellers and buyers to care whether an obligation was for sterling or not.

paper money unnecessary. The evidence we present, Grubb implies, is nothing more than the political posturing of the governor and does not accurately reflect the true state of affairs. That would have been an impressive rebuttal had we cited any statement by the governor, the proprietor, or anyone else directly engaged in these political machinations. We did not. Instead, we quoted a private letter from Pennsylvania's proprietary secretary to the proprietor, setting forth his opinion of the current state of affairs in Pennsylvania. We also quoted the Pennsylvania assembly and one of its committees. Those documents, the second and third sources in Appendix 1, make an impressive case and lend credibility to the other two.

In his rebuttal, Grubb (2006a, 59-60) makes one argument we find so curious as to be noteworthy:

> Penn knew that the temporary inflow of specie during the war would soon be gone and specie scarcity would return (which is consistent with Grubb's (2004, 340) new evidence series). In London on October 9, 1749 Penn wrote to his governor in Pennsylvania, '[E]very one is sensible that in two or three years almost the whole of the Gold and Silver that during the war was brought into the Colonys will be shipped hither, and wee shall have little but paper left . . . .' (Brock 1975, 356)

We find it difficult to reconcile Grubb's new data series with Penn's letter. According to Grubb's series, Pennsylvania possessed only 2,881 pounds of specie in 1749 when this letter was written—less than 3½ d. sterling per capita and barely 3.3 percent of the money supply. This was the great quantity of specie whose departure would return the colony to a condition of specie scarcity? So how does Penn's letter confirm Grubb's series?

Cross-colony circulation of bills of credit

Grubb (2004) concluded based on runaway ads that bills of credit of other colonies did not circulate in Pennsylvania, and Pennsylvania bills of credit did not circulate in neighboring colonies. This conclusion, as we pointed out (2006a, 12, fn. 6) is inconsistent with practically all the existing literature on the subject. Moreover, if Grubb is incorrect in his assertion, it invalidates his technique. In Michener and Wright (2006a, Appendix 2 and

fn. 8), we quote a large number of primary sources that explicitly describe the cross-colony circulation of bills. This is one instance that led Grubb to protest our "piling on" of "anecdotal quotes." Grubb questions the usefulness (2006a, 52) of such "anecdotal evidence . . . since there is tons of such on both sides of the debate." Therefore, it is fair to ask why Grubb did not select a more convincing "anecdotal quote" than Mazzei's to establish that bills of credit did not circulate in adjacent colonies. Grubb's use of Mazzei is discussed in Michener and Wright (2006a, 13) and defended by Grubb (2006a, 61). It is clear that Mazzei is referring to conditions post-1776 and (if one reads the complete passage) that Mazzei implies paper money sometimes circulated across state/colony boundaries. If this is the best Grubb can muster in support of his position, it belies his claim that "tons" of such evidence exist "on both sides of the debate."

Grubb objects to two bits of evidence we present pertaining to the circulation of Pennsylvania currency in Maryland circa 1760. One is a quote from Henry Callister, a Maryland merchant, who wrote in a private letter: "I said currency, which does not imply Maryland [paper] money, of which there is hardly any current—I think I was yet more particular, for I spoke of money and exchange as current in Pennsylvania which is our current money at present" (Michener and Wright 2006a, 25). Grubb accepts the authenticity of the quotation, but notes that Callister had unusually extensive dealings with Philadelphia merchants and was badly in need of cash. Those special circumstances, Grubb (2006a, 63-64) contends, account for Callister's willingness to accept Pennsylvania money. In our opinion, Grubb's counterargument is unsound. We quoted Callister describing what was and what was not circulating as money in his locale, not proclaiming his willingness to accept Pennsylvania currency. Grubb (2006a, 64) also writes that "In the rest of Callister's rather extensive correspondence with Wright, there is no indication that Pennsylvania currency was ever in frequent use in Maryland." Grubb takes no note of Callister's letter to a Mr. White, dated July 22, 1760 and cited in Michener and Wright (2006a, 25 fn. 16), reporting that "Pennsylvania and Jersey money . . . are current here."

The other bit of evidence is the Fitzhugh account books, also circa 1761-1764. Fitzhugh was a Maryland merchant, and the pivotal issue here is whether Fitzhugh kept his accounts in Maryland bills of credit during those years as Grubb (2004) maintains. If he did, then Grubb is correct that we would expect to see Pennsylvania currency being reduced to Maryland currency more often, if Pennsylvania currency had any substantial circulation within Maryland. We (2006a, 27) point out that during this period two parallel accounting systems were in use in Maryland and argue

that Fitzhugh was using the "hard currency" system described by McCusker (1978, 191), not Maryland bills of credit. Since the "hard currency" system used most of the same conventions as used in Pennsylvania, conversions to reduce sums of Pennsylvania money would rarely have been necessary and their absence would not be surprising. We (2006a, 26-27) noted that the value of £100 sterling differed substantially in the two systems, particularly in 1764, and that bills of exchange recorded in Fitzhugh's ledgers are priced at "hard currency" prices. Grubb (2006a, 63-64) reiterates his belief that Fitzhugh's ledgers are denominated in Maryland bills of credit, but does not address the exchange rate evidence we presented.

The implausible explosion of silver dollars

Our last objection (2006a, 16 and appendix 3) was to the large increase in the use of silver dollars that Grubb ostensibly finds between the end of the French and Indian War and 1775. We noted that in Pennsylvania silver dollars were substantially undervalued relative to gold Johannes and were, in fact, being exported during this period, consistent with Gresham's law. The overvaluation of Johannes, and the incentive it created to export all coins other than Johannes, was even remarked upon in contemporary newspapers. The reason for the increased use of "dollars" as a unit of account in runaway ads must lie elsewhere, we argued. Although our argument was based primarily on Gresham's law, Grubb did not address it. 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[hanne]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.

We went on to point out that there were at least three reasons, other than an increased use of Spanish silver, that might account for the increased use of dollars as a unit of account. In his rebuttal, Grubb heaps scorn on the third, which he characterizes (2006a, 65) as "the Maryland dollar gambit." "It is possible," we had written (2006a, 34), "that Pennsylvanians were led to make greater use of dollars as a unit of account by the more

extensive circulation of a dollar-denominated medium of exchange. If so, that medium of exchange was not silver but Maryland paper money." We noted that Maryland emitted paper money denominated in dollars beginning in 1767 and that there is evidence this paper money had at least a limited circulation in Pennsylvania. Grubb (2006a, 65) pounces on this statement. "If the switch to dollars by Pennsylvanians was caused by a flood of Maryland paper dollars into Pennsylvania so that they could be offered as rewards, as Michener and Wright imply, then Grubb's 2004 argument that rewards reflect media-of-exchange and unit-of-account money is upheld by Michener and Wright." We went out of our way to deny the implication. "The discussion above," we wrote (2006a, 37), "emphatically should not be taken to mean that every mention of a 'dollar' in late colonial Pennsylvania referred to Maryland's dollar denominated bills of credit, but only that, without further information, it *could* have referred to them. References to 'dollars' could also signify use of the dollar as a unit of account and hence simply be a means of accounting for a variety of other exchange media, including even *gold* coins." So we were not trying to have it "both ways" as Grubb charges. Our view is that there is a large component of custom in the unit of account, but that considerations of convenience can alter custom, albeit slowly. When Pennsylvania merchants raised the rated value of Johannes and half Johannes to £6 and £3 respectively, in 1767, they (perhaps inadvertently) made it easier to calculate in dollars, since the two coins now most favored by Gresham's law were assigned values that translated neatly into $16 and $8 respectively. Add to this the concomitant circulation of some of Maryland's new dollar-denominated paper money, and the convenience of using dollars as a unit of account increased. Over time, this greater convenience could undermine, to some extent, the pre-existing custom of using pounds as the unit of account. This is a plausible explanation for the explosion of 'dollars' in Pennsylvania runaway ads - more plausible, at least, than that Pennsylvanians were accumulating undervalued coins in violation of Gresham's law.

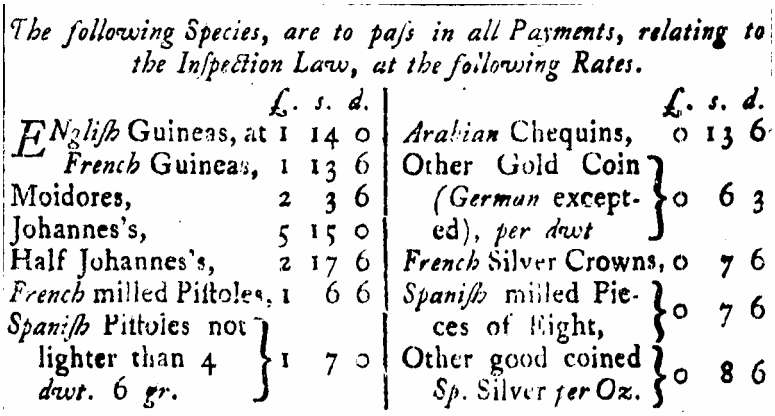
Grubb's (2006a, 66) comments on the geographic pattern of 'dollar' runaway ads—particularly his observation that Marylanders were less likely to place 'dollar' runaway ads than Pennsylvanians during the era of Maryland's dollar denominated currency—is new information to us and is worth pondering. It raises an interesting question: Why, if the circulation of dollar-denominated Maryland paper money in Pennsylvania influenced *Pennsylvanians* to adopt the dollar as a unit of account, did it not have a greater influence on *Marylanders*? The answer may simply be that most transactions in colonial America used neither specie nor paper money. The

overwhelming majority of everyday transactions were bookkeeping barter, where net balances were extinguished with such unconventional exchange media as labor and livestock. Bookkeeping barter was especially prevalent in rural areas; money (both paper and specie) was used infrequently outside major port cities. Where money was seldom used, how it was denominated would have had less influence on the customary unit of account. Marylanders might have been less influenced by their new dollar unit-of-account money because Marylanders encountered cash of all kinds less frequently than their brethren in Pennsylvania. However, suppose Grubb is correct on this point and Maryland's dollar money played no role in the increased use of dollars as a unit of account in Pennsylvania in the late colonial period. Our central point, which is that an increased use of Spanish silver dollars in colonial Pennsylvania violates Gresham's law and is thus implausible, remains untouched, as do the other two explanations we suggest to account for the phenomenon.

We submit, however, that Grubb's evidence (2006a, 66) that Maryland pounds remained the prevailing unit of account in Maryland during the "dollar-money" period—with "Maryland pounds" appearing in runaway ads as well as most merchant and government records—is extremely problematic for his own interpretation. "Maryland pounds" must be interpreted as a unit-of-account money, *because Maryland possessed no paper money denominated in Maryland pounds!* Maryland's pound-denominated paper currency was redeemed in 1764. The new issues, beginning in 1767, bore different denominations. Grubb notes that (2006a, 67) the new bills "reported on their face both a value in dollars and in pounds (Newman 1997, 167-169)," which is true, but misleading. The bills reported on their face both their value in dollars and their value in *pounds sterling*—that is to say, their redemption value, as the inscription on the bill makes unmistakably clear. The bill's value in *Maryland pounds* appears nowhere on the bill—nor were "Maryland pounds" the same as pounds sterling, as a casual glance at McCusker (1978, table 3.8) will quickly confirm. The reader can easily verify the tenor of the bills by examining sample bills from 1770 or 1774 on Notre Dame's web site: Link.

How, one may ask, were those bills assigned values in "Maryland pounds?" The definition of Maryland's unit of account prevailing in the late colonial period was that set by a tobacco inspection act passed in 1753, an act which included coin rating provisions (McCusker, 1978, 192). These ratings were published in almanacs, such as the *Maryland Almanack, for the year of our Lord, 1763*.

Figure 1: Maryland coin ratings



The following Species, are to pass in all Payments, relating to the Inspection Law, at the following Rates.

	£.	s.	d.		£.	s.	d.
English Guineas, at	1	14	0	*Arabian* Chequins,	0	13	6
French Guineas,	1	13	6	Other Gold Coin (*German* excepted), *per dwt*	0	6	3
Moidores,	2	3	6				
Johannes's,	5	15	0				
Half Johannes's,	2	17	6	*French* Silver Crowns,	0	7	6
French milled Pistoles,	1	6	6	*Spanish* milled Pieces of Eight,	0	7	6
Spanish Pistoles not lighter than 4 *dwt.* 6 *gr.*	1	7	0	Other good coined *Sp.* Silver *per Oz.*	0	8	6

Source: *Maryland Almanack for the year of our Lord, 1763*

One rated coin was the Spanish milled piece of eight (a.k.a "the dollar," see McCusker (1978, 7)), which was assigned a rating of 7½ shillings. Therefore, each dollar of Maryland's dollar-denominated paper currency was accounted as 7½ shillings in Maryland pounds.

REMARKS PERTAINING TO GRUBB'S USE OF THE HISTORICAL RECORD

Specie plenitude

One of Grubb's most fervent assertions is that we believe in "specie plenitude." We are accused of asserting that the colonies were (Grubb 2006a, 47) "awash in specie" and that the colonies (Grubb 2006a, 49) possessed "a large reservoir of specie . . . at all times." Moreover, many of Grubb's most blistering attacks center on his belief that we ignore or suppress information that contradicts us on this score. In one such attack Grubb (2006a, 47) notes that Brock (1975) "presents as many or more anecdotal quotes on specie scarcity as on specie plentitude. Michener and Wright simply ignore the specie-scarcity quotes."

We do not believe nor have we ever argued that specie was plentiful in colonial America. That we champion the belief that the colonies possessed a large reservoir of specie at all times is absurd: Michener (1987, 293-294) argues at length that colonial New England was entirely devoid of circulating specie for a generation and Wright (2001, 19-47; 2005, 44-65) devotes entire chapters to detailed discussions of the diversity of exchange media in early America! We *have* argued, though, that in many colonies and at many times there was as much or more specie in circulation than paper money. This position is easily confounded with another that we do not believe to be true, which is that specie was absolutely plentiful. There is no inconsistency in those views because in many colonies (Pennsylvania in particular) the quantity of paper money in circulation was often quite small.

Let us descend to particulars to free ourselves of this imputation. Jones's probate evidence provides an estimate of the total colonial money supply in 1774 that is roughly consistent with several literary estimates and also with McCallum (1992) (Michener 2003). We have reservations about the probate evidence (Michener and Wright 2006a, 18), but we have reluctantly embraced it. Jones's data indicate the total money supply for the Middle Colonies, reduced to its sterling equivalent, was £1.81 per capita. For the colonies as a whole, Jones (1980, 39, 128) finds a total money supply amounting to £1.087 sterling per capita. In 1774, there was approximately £0.58 sterling per capita in paper money circulating in the Middle Colonies, and the total supply of paper money in the colonies amounted to only about 3,000,000 dollars, or £0.287 sterling per capita (Michener 2003). The implication would appear to be that 68 percent of the money supply in the Middle Colonies and 74 percent of the money supply in the colonies as a whole consisted of specie. However, the implied quantities of specie—£1.23 sterling per capita in the Middle Colonies and £0.8 sterling per capita overall—are not large in any absolute sense. They compare quite unfavorably to estimates of the per-capita specie stock in England (£2.07-£2.15 sterling) or in France (£2.88 sterling) (McCusker and Menard 1988, 338, fn. 14).

In 1782, Mazzei, an Italian, made the following observation: "In 1773, the year disorders began, that is, ten years after the end of the previous war, all transactions were made almost entirely in specie, which, however, did not abound."[6] Grubb excoriates Michener (1988, 687) for using this quotation to establish that "all transactions were made almost

[6] Anyone interested in the full context of this quote can find it in Michener and Wright (2006a, fn. 13).

entirely in specie." "How," he asks rhetorically (2006a, 61), "did this Italian know so much about colony money supplies?" And how, he asks further, can this be when "Brock (1992) showed all the major colonies except Massachusetts had issued significant amounts of paper money after 1760 and/or had significant amounts outstanding in this period"? Finally, he accuses Michener (1988, 687) of truncating the quotation and leaving off the phrase "which, however, did not abound" so as to make it appear specie was plentiful.

Grubb is mistaken on all counts. "This Italian" (Mazzei) knew something about the medium of exchange in America because he lived in Virginia from 1773 to 1779 (Mazzei, Marchione, et. al. 1983 1: xxxviii). Far from being inconsistent with the quantity of paper money in Brock (1992), Mazzei's comment makes perfect sense as a statement about the colonies as a whole; Brock's estimates lead to the conclusion that the colonies as a whole possessed no more than £0.287 sterling per capita in paper money in 1774. Virginia, where Mazzei had first-hand experience, had less than £0.12 sterling per capita in paper money outstanding in December 1773, according to Brock (1992, 116). The quotation was truncated for terseness, nothing more. In fact, in a book, Wright (2001, 22) included the phrase "which, however, did not abound" because he had the luxury of space and because the clause is precisely consistent with our beliefs about specie in America on the eve of the Revolution.

In an apparent suggestion of legerdemain, Grubb sees villainy in our most innocuous statements. We (2006a, p. 5), in the midst of a *literature review* designed to demonstrate that the potential importance of Grubb's data series, remark that Grubb's data (if accepted) would have serious implications for those, such as McCusker and Menard, who argue the colonial money supply was adequate. That Grubb's data series has important implications for the adequacy of the colonial money supply is uncontroversial: Grubb says the same thing (Grubb 2004, 344). Later we (Michener and Wright 2006a, 15) contest the validity of Hamilton's money supply estimate, which is a linchpin of McCusker and Menard's position. In a lengthy discussion, Grubb (2006a, 51) derides our "opening contradiction" which is that we "[appeal] to McCusker and Menard (1985, 338) as support for their position," (Grubb's misapprehension that we believe in specie plenitude leads him to assume we agree with McCusker and Menard) "and then [trash] the evidence that generated that support."

Probate Evidence

The debate about the correct interpretation of the probate evidence is lengthy, beginning with Michener (1987, 275-276), then passing to Grubb (2004, 342-343), Michener and Wright (2006a, 15-19) and Grubb (2006a, 60-61). At the end (Grubb 2006a, 60-61), Grubb's argument is in ruins, and he responds with innuendo. One example is Grubb's contention that Michener (1987) is somehow responsible for the brouhaha because he was the first to use Jones's probate evidence to estimate the specie stock, and that, though it was not "easily done," "Michener manipulated Jones's evidence to make it fit his view." But Weiss (1970, 779) was the first to use probate evidence in this way, and both Weiss (1970) and Michener (1987) proceed in exactly the same fashion—by subtracting the known quantity of paper money from Jones's estimate of the total money supply to derive an estimate of the specie portion of the money supply. There is nothing either difficult or devious about this approach.[7] Grubb goes on to dismiss our refutation of his own calculations as nothing more than a "lengthy diatribe of remanipulation" that validates his original point, which was "that the Jones evidence can be plausibly manipulated to say almost anything." To this, we can only reply that there is scant support in Grubb (2004, 342-343) for the notion that this was Grubb's original point and that Michener and Wright (2006a, 15-19) demonstrate convincingly that Grubb's calculations are insupportable.

Grubb's attack on the Fixed Exchange rate thesis

The entire question of fixed exchange rates is of little relevance to the issue of whether Grubb's data series is accurate. Everyone, including Grubb, agrees that paper monies issued in the Middle Colonies maintained a stable value in the late colonial period. The debate centers on why. In Michener (1987, 1988, 2003) and Michener and Wright (2005), we have presented a theory to account for "why," but our criticism of Grubb

[7] Roger Weiss (1970, 779) arrived at a smaller fraction of specie in the money supply. There is a simple explanation: Weiss, whose article was published in 1970, based his analysis on Jones's 1968 dissertation rather than her 1980 book. In her dissertation, Jones (1968, Tables 3&4, 50-51) estimated the money supply in the three Middle Colonies at £2.0 local currency per free white capita. After converting to sterling, Weiss began with an estimated total money supply of £1.2 sterling per capita, rather than Jones's more recent estimate of £1.81 sterling per capita.

(2006a) in no way hinges on our being correct on this issue. Grubb tries to argue that if colonial paper monies did not exchange at exactly fixed rates with one another, they could not have circulated across colony borders, so that accepting our "fixed rate" thesis is a prerequisite for accepting our critique of Grubb. However, belief in the cross-colony circulation of bills of credit would hardly be limited to subscribers of our fixed-rates theory; many economic historians who accept the cross-colony circulation of bills of credit reserve judgment on it and Smith (1988) (to mention a noteworthy example) believes in the cross-colony circulation of bills while vigorously disputing our theory. Moreover, Grubb's contention that the cross-colony circulation of bills could occur only where relative values of bills were absolutely fixed is doubtful (though, where they were fixed in value, it no doubt facilitated that cross-colony circulation). It is well known that in colonial times many book debts were extinguished with such variegated media as chickens, livestock, and day labor—media far more heterogeneous and variable in their value than, say, New Jersey bills of credit.

Moreover, Grubb's attack on the fixed exchange rate thesis is inaccurate as well as irrelevant. We cite four particulars.

> 1) Grubb says (2006a, 54): "Michener and Wright assume that exchange rates were universally fixed in the colonial period." We make no such claim. In Michener (1987, 288-294) and Michener and Wright (2006b, 26, 29-30) one of the central messages is that in colonial New England all the circulating specie had been exported by about 1713 and that exchange rates floated until the Massachusetts currency reform in 1750. That was why New England's paper money depreciated even though the paper money in most other colonies did not.
>
> 2) Grubb cites McCusker's exchange rate data (2006a, 55-56) to show there was "no fixity or constancy in the exchange rate." As we have repeatedly stressed (see, for example, Michener 1987, 265-266; Michener and Wright 2006b, 26-27, 31-32) our proposition is that coin rating fixed the *par of exchange* and that market exchange rates fluctuated about par within specie points, which were relatively wide in the 18th century because of higher transportation costs. The genuine issue is whether the variability of exchange rates documented by McCusker is small enough to be consistent with fluctuations about par within specie points. To that end, Michener (1987, 265-275) presents several histograms of exchange rate fluctuations and

argues that, in many instances, the fluctuations can be so explained. Therefore, simply noting that McCusker's exchange rates are not constant hardly "destroys Michener and Wright's core model of the colonial monetary system," as Grubb (2006a, 55) contends.

3) Grubb (2006a, 55) accurately attributes to us the view that the coin rating agreements that lie at the heart of our theory were "created, maintained and enforced by custom . . . [and] by agreements among merchants in the marketplace." However, he contends that we "present no direct evidence of such a merchant cartel." That is incorrect: Michener and Wright (2005, 2006b, 24), Michener (2003, fn. 19), and Michener (1987, fn. 25) do present "direct evidence" of the merchant agreements underlying the effectiveness of coin rating. One bit of evidence comes from almanacs. Grubb (2006a, 55) objects that the tables adduced tell us no more about the fixity of coin ratings than the presence of a currency exchange table in the *Wall Street Journal*—an objection that misses the point that coin ratings were published not in colonial newspapers, but in almanacs, which then, as now, were annual publications.[8]

4) To discredit the model in Michener (1987), Grubb (2006a, 47) caricatures it. That model assumes, Grubb (2006a, 49) asserts, that:

i) There is a perfectly fixed exchange rate between a colony's paper money and foreign specie monies (they are perfect substitutes),

ii) money demand is invariant over time even in the short-run,

iii) there is a large reservoir of specie in the colony at all times, and

iv) transaction and information costs are zero.

Michener (1987, 253-256) does introduce a simple formal model as an expository device that incorporates some of those elements. However, Michener holds that the first condition was only met in some colonies at some times, and even in those colonies, condition (iii) was unnecessary. The stock of circulating specie only had to be positive. Moreover, Michener (1987, 256-257, 265-266, 277, 283) explicitly acknowledges the oversimplifications implicit in the model

[8] Grubb (2006a, 55) also objects that "these almanac coin-rating tables are for unit-of-account exchange rates and not media-of-exchange rates (Michener and Wright are hoisted on their own petard here)." This objection baffles us; the distinction between "unit of account exchange rates" and "medium-of-exchange rates" seems little more than word play.

and distances himself from them. An excellent example is transactions costs; Michener (1987, 266) explicitly discusses specie points. Moreover, Michener (1988, 691) and Michener (2003) explicitly discuss factors that influence money demand: War, Michener argues, disrupted bookkeeping barter by uprooting people from their communities and placing them in mortal peril, increasing money demand. Grubb's summary of our position is neither accurate nor fair-minded.

We could go on, but won't, as we believe that at this point any scholar who has followed the debate closely and objectively will harbor deep reservations about Grubb's monetary history papers.

REFERENCES

Brock, L. V. 1975. *The Currency of the American Colonies, 1700-1764: A Study in Colonial Finance and Imperial Relations*. New York: Arno Press.

Brock, L. V. 1992. Colonial Currency, Prices & Exchange Rates. *Essays in History* 34: 74-132.

Carey, M. 1786. *Debates and proceedings of the General Assembly of Pennsylvania on the memorials praying a repeal or suspension of the law annulling the charter of the bank*. Philadelphia: Printed for Carey and Company.

Grubb, F. W. 2001. The Circulating Medium of Exchange in Colonial Pennsylvania, 1729-1775: New Estimates of Monetary Composition and Economic Growth. Working Paper, University of Delaware Department of Economics.

Grubb, F. W. 2003. Creating the US Dollar Currency Union, 1748-1811: A Quest for Monetary Stability or a Usurpation of State Sovereignty for Personal Gain? *American Economic Review* 93(5): 1778-1798.

Grubb, F. W. 2004. The Circulating Medium of Exchange in Colonial Pennsylvania, 1729-1775: New Estimates of Monetary Composition, Performance, and Economic Growth. *Explorations in Economic History* 41(4): 329-360.

Grubb, F. W. 2005. Creating the U.S. Dollar Currency Union: A Reply to Michener and Wright — Including a New View from Canada. *American Economic Review* 95(4): 1341-1348.

Grubb, F. W. 2006a. Theory, Evidence, and Belief—The Colonial Money Puzzle Revisited: Reply to Michener and Wright. *Econ Journal Watch* 3(1): 45-72. Link.

Grubb, F. W. 2006b. The Constitutional Creation of a Common Currency in the U.S., 1748-1811: Monetary Stabilization Versus Merchant Rent Seeking in Jurgen Nautz and Lars Jonung, eds., *Conflict Potentials in Monetary Unions*. Stuttgart: Steiner Verlag, forthcoming.

Hepburn, A. B. 1967. *A history of currency in the United States*. 1903. Reprint. New York: Augustus M. Kelley.

Hockley, R. 1753. Richard Hockley to T. Penn, 24 February 1753. Penn Official Correspondence, 1747-1771, HSP vol. 6: 17.

Hutson, J. 1970. Campaign for a Royal Province. *Pennsylvania Magazine of History and Biography* 94(4): 427-463.

Jones, A. H. 1968. Wealth Estimates for the American Middle Colonies, 1774, Ph.D. diss., University of Chicago.

Jones, A. H. 1980. *Wealth of a Nation to Be: The American Colonies on the Eve of the Revolution*. New York: Columbia University Press.

Maryland Almanack, for the year of our Lord, 1763. 1762. Annapolis: Printed by Jonas Green.

Mazzei, F., M. Marchione, et al. 1983. *Philip Mazzei: Selected Writings and Correspondence*. Prato, Italy: Cassa di Risparmi e Depositi di Prato.

McCusker, J. J. 1978. *Money and Exchange in Europe and America, 1600-1775*. Chapel Hill: University of North Carolina Press.

McCusker, J. J. and R. R. Menard. 1985. *The Economy of British America, 1607-1789*. Chapel Hill: North Carolina, University of North Carolina Press.

Michener, R. W. 1987. Fixed Exchange Rates and the Quantity Theory in Colonial America. *Carnegie-Rochester Conference Series on Public Policy* 27: 233-307.

Michener, R. W. 1988. Backing Theories and the Currencies of Eighteenth-Century America: A Comment. *Journal of Economic History* 48(3): 682-692.

Michener, R. W. 2003. Money in the American Colonies. EH.Net Encyclopedia. R. Whaples. Link.

Michener, R. W. and R. E. Wright. 2005. State "Currencies" and the Transition to the U.S. Dollar: Clarifying Some Confusions. *American Economic Review* 95(3): 682-703.

Michener, R. W. and R. E. Wright. 2006a. Miscounting Money of Colonial America. *Econ Journal Watch* 3(1): 4-44. Link.

Michener, R. W. and R. E. Wright. 2006b. Development of the U.S. Monetary Union. *Financial History Review* 13(1): 19-41

Morris, R., E. J. Ferguson, et al. 1973. *The Papers of Robert Morris, 1781 - 1784*. Pittsburgh: University of Pittsburgh Press.

Pennsylvania Evening Post. Philadelphia, Pa: B. Towne.

Pennsylvania Packet, and the General Advertiser. Philadelphia, Pa.: J. Dunlap

Smith, B. D. 1988. The Relationship Between Money and Prices: Some Historical Evidence Reconsidered. *Quarterly Review* 12(3): 18-32.

Tilghman, J. 1768. Letter to Sir William Johnson, 29 November 1768. Printed in *The Papers of Sir William Johnson, 1715-1774*. Albany: The University of the State of New York 6: 507-508.

Weiss, R. W. 1970. The Issue of Paper Money in the American Colonies, 1720-1774. *Journal of Economic History* 30(4): 770-784.

Wright, R. E. 2001. *Origins of Commercial Banking in America, 1750-1800*. Lanham, Md.: Rowman and Littlefield.

Wright, R. E. 2005. *The First Wall Street: Chestnut Street, Philadelphia, and the Birth of American Finance*. Chicago: University of Chicago Press.

ABOUT THE AUTHORS

Ronald W. Michener is associate professor of economics at the University of Virginia and can be reached via e-mail at rwm3n@virginia.edu. During the last 30 years he has assembled a massive collection of primary source materials related to America's colonial monetary and political history.

Robert E. Wright is clinical associate professor of economics at New York University's Stern School of Business. He can be reached via e-mail at rwright@stern.nyu.edu. Over the last 10 years, he has read thousands of manuscripts and other primary source materials related to colonial Pennsylvania's business, economic, and political history.

Econ Journal Watch,
Volume 3, Number 2,
May 2006, pp 275-282.

Economics in Practice: Follow Up

The Costs of Critical Commentary in Economics Journals

ROBERT WHAPLES *

THE BENEFITS OF CRITICAL COMMENTARY ARE MANIFEST. Indeed, all of human understanding depends upon it. Coelho, De Worken-Eley, and McClure (2005) document that critical commentary declined as a share of the pages published in five highly-ranked economics journals between 1963 and 2004. They argue that this decline constitutes a negative trend, chastising journal editors for this mistake, while enumerating several benefits that arise from commentary—especially the discovery and advertisement of errors and limitations, but also allowing readers and researchers to achieve a broader and deeper comprehension, constraining editors' self-serving behavior, and piquing readers' interest. They argue that "an editorial posture that eschews critical commentary subjugates the spirit of scientific inquiry," and suggest that editors' ignorance of the benefits are at the root the problem (360).

Unfortunately, however, commentary has costs as well as benefits so its optimal amount is finite.

The art of editing a journal is an exercise in constrained optimization. Editors' decisions are immensely complex, but one version of the optimization exercise would be to maximize the quality of the work published in the journal or the journal's prestige subject to the space allowed, quantity and quality of the articles submitted, hassles involved in editing, and other goals. This perspective offers a broader interpretation of the decline in commentary that Coelho, De Worken-Eley, and McClure identify.

* Department of Economics, Wake Forest University.

During the 1970s institutional and technological changes began to alter the calculus of journal editors regarding commentary. Perhaps foremost was the launch of the *Social Sciences Citation Index* (SSCI) in 1973. This index allowed interested parties to easily quantify the success of a published article by tallying up the number of times the article was cited. By the early 1980s articles began to appear that used the *SSCI* to rank economics departments and journals themselves (e.g. Davis and Papanek 1984; Liebowitz and Palmer 1984). For better or for worse, the *SSCI* became widely accepted as a useful gauge for publishing success.[1] Shortly thereafter, as Coehlo et al. verify, some journal editors announced a policy shift away from publishing commentary.

Coelho et al. lament that the editors of the *American Economic Review* (*AER*) found their own "prejudice against critical commentary" to be desirable. Perhaps this policy wasn't so much a *prejudice*—i.e. a groundless partiality—but a decision that reflected an implicit or explicit weighing of the perceived benefits versus the perceived costs of publishing commentary.

To assess the prudence of the policy one can quantitatively compare the citation success of comments to that of standard journal articles published during the early and mid 1980s, the period preceding the decision of the *AER* (and implicitly of other top journals as well) to change its editorial policy.

From the vantage point of the present, commentary published during this period was much less likely to be cited than standard articles. Tables 1, 2, and 3 compare the citations per page published in comments and standard articles in three of the journals Coelho et al. examine: *American Economic Review*, *Journal of Political Economy* (*JPE*), and *Quarterly Journal of Economics* (*QJE*).[2] I use two definitions of commentary. The first version (Commentary I) defines as commentary only articles that bear the title "Comment" or "Reply." The second version (Commentary II) is broader including articles in the *JPE* referred to as "Confirmations and Contradictions" plus articles in the *AER* and *QJE* that are explicit comments upon other articles, but which aren't labeled as such.[3]

---

[1] Indeed, the publisher of the *SSCI* eventually began touting its indices' usefulness in selecting Nobel Prize winners (Yancey 2005). Likewise, many journals now advertise their "impact factor" and ranking using information supplied by the publisher of the *SSCI* .

[2] Citation counts were calculated during the week of December 19, 2005. The relationship between article length and citations appears to be linear. In a regression of citations versus length the quadratic term is statistically insignificant.

[3] All calculations in Tables 2 and 3 exclude a small subset of articles that are neither commentary nor standard articles. These include "miscellany" articles in the *JPE* and articles in the *JPE* and *QJE* in appreciation of the careers of eminent economists. The *Social Sciences*

Table 1
Citations per Page: Commentary vs. Standard Articles in the *American Economic Review* (1980-1985)

	Commentary I N=215	Commentary II N=225	Standard Articles N=526
Citations per page	1.77	2.48	4.71
75th percentile	2.00	2.00	5.13
50th percentile	0.82	0.83	2.27
25th percentile	0.00	0.00	1.00

Table 2
Citations per Page: Commentary vs. Standard Articles in the *Journal of Political Economy* (1983-1986)

	Commentary I N=21	Commentary II N=29	Standard Articles N=215
Citations per page	1.37	1.41	3.19
75th percentile	1.83	1.97	3.41
50th percentile	0.52	0.57	1.48
25th percentile	0.00	0.00	0.75

Citation Index's coverage of *JPE* articles before 1983 is fairly incomplete, so I ignored these years and added in data from 1986.

Table 3
Citations per Page: Commentary vs. Standard Articles in the *Quarterly Journal of Economics* (1980-1985)

	Commentary I N=38	Commentary II N=46	Standard Articles N=244
Citations per page	0.41	0.44	1.63
75th percentile	0.45	0.50	1.56
50th percentile	0.14	0.18	0.54
25th percentile	0.00	0.00	0.22

Clearly, standard articles published during this period have subsequently been cited much more frequently. Across the three journals, citations per page of commentary (narrowly defined) ranges from only 25 percent to 43 percent of that for standard articles. This result is not driven by a few outliers, as the entire distribution for standard articles is considerably higher.[4]

Regression results show that these differences are robust and statistically significant. The nine regressions whose coefficients are reported in Table 4 control for number of authors, year of publication, and topic and type of article. The last variant in each case omits the top and bottom five percent of articles as measured by citations per page in case a few outliers are driving the results. Apparently, they are not.[5] (The complete results for the nine regressions reported in Table 4 are contained in an Appendix linked at the end of this article.)

[4] The *AER*'s citation average for Commentary II is much higher than for Commentary I due to a single outlier.

[5] Sadly, from my point of view, in every case the coefficients for economic history articles and history of economics articles are negative—and these are often statistically significant. This may help explain why top journals publish such articles infrequently. This may also reflect problems with the *SSCI* as argued by Weintraub (2006).

Table 4
Impact of Commentary on Number of Citations per Page

	AER	*JPE*	*QJE*
Commentary I	-3.145*	-1.548	-1.098**
	(0.737)	(1.108)	(0.570)
Commentary II	-2.204*	-1.730**	-1.134*
	(0.732)	(0.958)	(0.523)
Commentary II	-2.275*	-0.928*	-0.506*
Outliers omitted	(0.244)	(0.395)	(0.177)

Notes: Dependent variable = citations per page. Independent variables not reported in the table are number of authors, year of publication, and dummies for miscellaneous articles, appreciation articles, and articles concerning the history of economic thought and economic history. Standard errors are in parentheses.
* = statistically significant at the 95% confidence level.
** = statistically significant at the 90% confidence level.

The regression results are fairly consistent with the averages reported in Tables 1, 2, and 3. Eyeballing contemporaneous data found in the *Social Sciences Citations Index* an editor at a top journal would likely have noted the same phenomenon—that commentary doesn't pay off for these journals in terms of citations, often cutting citations per page in half. Hence, commentary is costly. The opportunity cost of publishing a comment (and the almost inevitable reply) is that the journal cannot publish as many standard articles. Standard articles are cited more frequently, and hence probably are more effective in building and maintaining the prestige of the journal.

The results above cannot precisely identify the number of citations lost from publishing the *marginal* article (the one on the bubble between being accepted and rejected) rather than a few comments and replies that take up the same amount of journal space. However, they may come close, especially the regression that throws out superstar and dud articles.[6] In

[6] One referee suggested that if comments get more citations than the articles that are commented upon, this would suggest that too few comments are being published. In fact, these comments receive far fewer citations than commented upon articles, whose citation levels are somewhat higher than those of all standard articles. It is possible that published comments increase the number of citations to the articles that they comment upon, but also possible that widely-cited articles attract more comment, so these differences are difficult to interpret. In theory one could test to see whether published comments have an impact on total citations to an article. The best test might be to compare articles that are otherwise similar, all of which attract comments of equal importance and quality, but only some of

addition, these results need not hold for all economics journals—especially lower-ranked journals and subfield journals. An earlier study (Whaples 2002), finds that the citation rates for comments and replies in the *Journal of Economic History* do not differ from standard articles.

The discussion up to this point has ignored some of the additional costs of publishing commentary. There are fixed costs of handling each submission to a journal that are the same regardless of the length or type of submission. In fact, the fixed costs of comments may be higher than for standard articles because they involve additional correspondence and negotiation with the original author. In addition, there may be psychic costs to the editor from publishing commentary—the implicit recognition that the editor failed to spot some of the article's flaws.

My argument does *not* state that the decline in critical commentary is a good thing. It points out an important cost of commentary and a possible motivation for the trend documented by Coelho et al.—why a set of producers decided to drop or downsize a product line. Nor does my argument suggest that measuring the success of an article or journal by volume of citations is wise. There are both costs and benefits of using citations and the *Social Sciences Citation Index* has been sharply and effectively criticized because of its opaque, circular, and even ideologically-biased criteria for selecting journals to include in the index (Klein and Chiang 2004). Moreover, while Coelho et al.'s data verify the decline in critical commentary in the top journals, they do not show that critical commentary more broadly has declined. Perhaps the decisions of these top journal editors pushed commentary into other channels. Perhaps they were also partly in response to the proliferation of other outlets, such as the launching

which have the comment published. Performing such a comparison seems unlikely, however.

The table below gives the distribution of citations for the articles on which comments were published between 1980 and 1985. The *JPE* is omitted because the *SSCI*'s coverage of it before 1983 is too sporadic. The calculations below subtract the comments themselves from the articles' citation totals.

	AER N=58	QJE N=17
75th percentile	8.71	2.22
50th percentile	4.46	0.75
25th percentile	1.58	0.21

of *Economic Letters* in 1978. Finally, the decline in commentary may also have been partly due to a decline in the supply, as changes in tastes or technology increased the desire and ability to produce standard articles in comparison to commentary.

Innovations which reduce the costs of commentary are surely a good thing, but it is imperative that one does not ignore the costs of commentary altogether.

APPENDIX

This Excel file contains the nine regressions from which the coefficients in Table 4 are taken.

REFERENCES

Coelho, Philip R.P., Frederick De Worken-Eley III, and James E. McClure. 2005. Decline in Critical Commentary, 1963-2004. *Econ Journal Watch* 2(2): 355-361. Link.

Davis, Paul and Gustav F. Papanek. 1984. Faculty Ratings of Major Economics Departments by Citations. *American Economic Review*. 74(1): 225-230.

Klein, Daniel B. and Eric Chiang. 2004. The Social Science Citation Index: A Black Box—with an Ideological Bias? *Econ Journal Watch* 1(1): 134-65. Link.

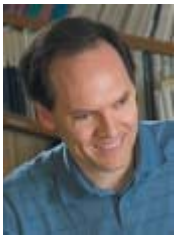
Liebowitz, S.J. and J.P. Palmer. 1984. Assessing the Relative Impacts of Economics Journals. *Journal of Economic Literature*. 22(1): 77-88.

Weintraub, E. Roy. 2006. Correspondence. *Econ Journal Watch*. 3(1): 206-209. Link.

Whaples, Robert. 2002. The Supply and Demand of Economic History: Recent Trends in the Journal of Economic History. *Journal of Economic History*. 62(2): 524-532.

Yancey, Rodney. 2005. Nobel Predictions: Introducing the 2005 Thomson Scientific Laureates. Link.

ABOUT THE AUTHOR



Robert Whaples is Professor of Economis at Wake Forest University. He is Director of EH.NET, which is dedicated to making economic history accessible to the public (at www.eh.net) and editor of EH.NET's Encyclopedia of Economic and Business History. Among his recent papers is "Time to Eliminate the Penny from the U.S. Coinage System: A Critical Analysis." Email: whaples@wfu.edu.

GO TO COEHLO ET AL. (2005) ARTICLE

Econ Journal Watch,
Volume 3, Number 2,
May 2006, pp 283-291.

ECONOMICS IN PRACTICE: FOLLOW-UP

Why Has Critical Commentary Been Curtailed at Top Economics Journals? A Reply to Robert Whaples

PHILIP R.P. COELHO AND JAMES E. MCCLURE *

IN OUR ARTICLE WITH FREDERICK DE WORKEN-ELEY III IN THE April 2005 issue of this *Journal*, we documented the decline in critical commentary (i.e., comments, replies, rejoinders) that occurred between 1963 and 2004 in the top general interest journals in economics. Explaining the decline was not our focus, although we lamented the decline because it makes the journals less valuable as forums for discussion.

Robert Whaples offers an explanation for the decline and its timing: Journal editors seek to advance their journal's reputation, and with the advent in 1973 of the *Social Sciences Citation Index*, reputation became closely identified with measured citations, and as Whaples shows, regular articles receive more citations per page than do critical commentary pieces. Editors displaced commentary with more regular articles to enhance citations to their journal.

Whaples' empirical analysis employs a narrow and a broad measure of critical commentary and finds t-statistics that indicate that cites per page are negative and highly significant statistically. The t-statistics, however, do

* Coelho and McClure: Department of Economics, Ball State University. Muncie, IN 47306.
The authors acknowledge the generous assistance provided on earlier drafts by David Laband, Judy Lane, Tung Liu, Michael Piette, Lee Spector, and Robert Whaples.

not tell the whole story. As Ziliak and McCloskey (2004) explain, statistical significance can be achieved with a large enough sample size. For this and other reasons, they are vehement about the over-reliance upon the t-statistics on coefficient estimates. They emphasize the crucial importance of a broad perspective and interpretation. To assess whether the decline in commentary was driven by concerns for citations, the explanatory power of his model, the sample size, the statistical significance and size of coefficients are all crucial considerations. Although Whaples discusses coefficient size and significance, he fails to interpret them in the context of the relatively large sample size and the low Adjusted R-squared statistics in his model. The data in Table 1 summarize these statistics from Whaples' appendix:

Table 1

Adjusted R-Squared Statistics from Whaples' appendix: (by measure of commentary and by journal) [1]			
	AER	*JPE*	*QJE*
Commentary I *(narrow measure)*	.023	.029	.004
Commentary II *(broad measure)*	.011	.034	.008

The table shows that in Whaples' regressions (again, where citations per page is the dependent variable with alternative measures of commentary as an independent variable), the percentage of the variation explained is 2, 3, and 1 percent for the *AER*, *JPE*, and *QJE* respectively (rounded to the nearest percentage point), using the specification that gives the highest adjusted R-squared statistics. Doing the obvious arithmetic, the corresponding unexplained variations in the regressions on citations per page (in percentages) are: 98, 97, and 99. The magnitude of the unexplained

[1] Whaples' "outliers omitted" regressions are excluded from consideration. Whaples provides no reason for excluding the top and bottom 5 percent of citation-per-page-articles beyond identifying them as the top and bottom 5 percent. Whaples suggests that editors focused on increasing citations per page. If articles that are "superstars" or "duds" in terms of citations per page impact journal reputations, their inclusion is necessary for an assessment of Whaples' hypothesis. For an interesting exchange about the relevance of articles that have received no citations, "dry holes", in the assessment of research productivity, see: Laband and Tollison (2003), Mayer (2004), and Laband and Tollison (2004). Omitting, "superstar" articles, on the other hand, is likely to influence the measured "impact" of the journal; Brauninger's and Haucap's (2003) show the total effect of impact upon reputation provides evidence against omitting outliers from assessments regarding journal reputations.

variation in Whaples' model suggests that alternative specifications or entirely different models may be superior depictions and predictors of economic reality.

Still, Whaples was on to something and led us to think about why the decline in critical commentary occurred when it did. We agree with Whaples that citation seeking probably had something to do with it, but we ask: citation to what? Whaples focuses on citations to the journal. We propose that citations *to editors and editorial board members* provide a motivation to alter editorial policies. Economists such as Adam Smith, James Duesenberry, and Robert Frank have explored the desire for esteem, status, and prestige in social affairs, and philosophers of science emphasize the role of esteem in scholarship. Every economist knows intuitively that most economists desire esteem, recognition, and prestige, as means to other ends, and as ends in themselves.

David Laband, Robert Tollison, and Gokhan Karahan (2002) analyzed the editorial screening process at top journals in economics. They examined the citations to editors and members of the editorial board of the *American Economic Review* by regular articles appearing in the *AER*. They compared that to the number of citations to *AER* insiders by regular articles appearing in the *Journal of Political Economy* and the *Quarterly Journal of Economics*. Their results (326) are reproduced below in Table 2.

Table 2

	1985	1990	1995	2000
AER references to *AER* editors and editorial board members, per article	0.396	0.727	1.444	3.022
JPE/*QJE* references to *AER* editors and editorial board members, per article	0.420	0.522	0.826	0.761

As Laband *et al.* point out (326), the divergence by the year 2000 is truly remarkable: regular articles in the *AER* cite *AER* insiders at a rate *four times* that these insiders are cited in regular articles published in the *JPE* and *QJE*.

Still, Laband *et al.* investigated only regular articles, excluding critical commentary. We extend their method to critical commentary in the *AER*,

and investigated only the years 1985, 1990, 1995 and 2000.[2] Citations per critical commentary article to *AER* insiders are shown in Table 3:[3]

Table 3

	1985	1990	1995	2000
AER references to *AER* editors and editorial board members per critical commentary article	0	.091	1.10	.083

Authors of comments, replies, and rejoinders build on the scholarship of the original article and introduce few additional references. The tabular results confirm the expectation that the citation "payoff" to editorial insiders of critical commentary pales in comparison to the citations to editorial insiders generated by regular articles. If the quest for citations to their own work motivates editorial insiders, they would be expected to be biased against the publication of critical commentary, and in favor of the publication of original articles.

Laband *et al.* estimate that the extra citations associated with an editorial position creates a "salary premium" for editors and editorial board members "on the order of 10-20 percent in base salary" (326). They suggest that in the case of a 10 percent premium the lifetime present value of additional references to a typical editor or board member is likely to be on the order of $240,000.[4] Because these estimates are for the gains associated with an editorial position, the payoff from the reduction in

[2] We contacted Laband and he generously provided us the lists of names of editors and editorial board members that were used in his analysis for 1985 and 2000; following his procedure of using the previous year staff, we compiled the 1990 and 1995 years ourselves. We then searched for these names in reference lists of comments, replies and rejoinders. The lists of names used are appended to this paper following the references.

[3] In 1985 there were zero citations to insiders; in 1990 there were two citations to insiders in 22 critical commentary articles; in 1995 there were eleven citations to insiders in 10 articles; in 2000 there was one citation to an insider in 12 articles.

[4] They give an even lower bound estimate under the assumption that the salary premium for additional citations due to an editorial position is only 5 percent; in this case the value of the editorial position (additional references) is $120,000. Via email correspondence, Laband confirmed that $120,000 was the correct value as opposed to the misprint of $20,000 that appeared on page 327 in Laband *et al.* The present value of the upper-bound premium (not reported by Laband *et al.*) is approximately $500,000 (in the case where the salary premium is 20 percent).

commentary (a shift in an editorial policy) is likely to be substantially smaller.

Laband *et al.* downplayed the importance of the pecuniary incentives they uncovered because well established economists already receive substantial incomes.[5] But additional citations to editorial insiders enhance not only their money income, but also their relative standing or prestige in the profession.

Why did editors sharply reduce critical commentary? Whaples suggests that with the advent of the SSCI in the mid-1970s, editors replaced commentary pieces with more regular articles because they wanted to enhance the citation-measured reputations of their journals. We suggest that editors reduced critical commentary partly because regular articles offered them more citations to themselves, providing them marginally higher incomes and higher prestige. Both Whaples' and our hypotheses predict a decline in critical commentary, but a tension is possible. If *AER* editorial decision-making is affected by the number of citations to editorial insiders, and articles citing *AER* insiders receive fewer citations in other journals than articles with few citations to insiders, then the net result may be fewer citations to the *AER*. It would be interesting to compare the two following cites per page in non-*AER* journals: 1) cites to *AER* articles that cite *AER* insiders often, and 2) cites to *AER* articles that cite *AER* insiders infrequently.

If, as Whaples suggests, editors were motivated by concerns for their journals' reputations, we question whether the curtailment of commentary did in fact increase *SSCI* citations per page. *Alternative* citation-based measures (alternatives to citations per page) of journal quality, provided by Laband and Piette (1994a), suggest that the top journals experienced declining reputations over the twenty year period from 1970 to 1990. They show that between 1970 and 1990, the "percentage of impacted adjusted citations" fell for: the *AER* (from 19.4 to 11.9 percent); the *JPE* (from 11.8 to 9.3 percent); the *QJE* (from 8.0 to 4.7 percent); the *EJ* (from 3.9 to 1.2 percent) (Leband and Piette 1994a, 654). These four journals combined had 44.1 percent of impact-adjusted citations in 1970, this fell to 26.1 percent of impact-adjusted citations in 1990 (unadjusted citations fell from 29.8 percent to 16.6 percent for these four journals). In terms of their

[5] We are less charitable (or more jaundiced) than Laband *et al.* on this score: we believe that the money involved with an editorial position (reasonable estimates from present values of from $240,000 to $500,000) present substantial moral hazards that could easily affect editorial decisions.

percentages of citations, these journals experienced marked declines in their reputations. Other factors, such as the explosion in the number of journals, may have played a role, but these basic results should raise doubts about the effectiveness of editorial policy at the top journals.

Another concern bearing on the wisdom of editorial prejudice regarding commentary is the existence of editorial "favoritism" investigated by Laband and Piette (1994b); by favoritism they mean that editors are "publishing substandard material written by friends/allies" (202). In their conclusion Laband and Piette speculate that the costs of such publications could be justified *if* journals obtain benefits by being able to get higher quality editors than otherwise. Laband and Piette (1994b) argue that empowering editors with the ability to extend favors can sensibly be viewed as a form of compensation because favors tend to be reciprocated by friends/allies. To our knowledge, these costs and benefits remain unquantified leaving the issue of whether editorial favoritism is justified unresolved. Additionally, Laband's and Piette's (1994b) revelation of editorial favoritism raises the following issue: If an editor allows a friend/ally to publish a substandard paper, would he want to publish critical commentary that exposes errors in the journal?

CONCLUSION

Robert Whaples states: "The benefits of critical commentary are manifest. Indeed, all of human understanding depends upon it." We concur—commentary is crucial to human understanding, economics is not an exception. Whaples suggests that the decline of commentary and its timing are a result of editors who: 1) responded to the *SSCI* measures of citations in the social sciences by realizing that citations-per-page were now the leading metric of journal quality; and, consequently, 2) curtailed publications of commentary to make room for more original articles because their citation-per-page exceeded that of commentary. The sizes of the pertinent Adjusted R-squared statistics characterizing Whaples' regressions and other factors raise concerns that the motivating factor that Whaples proposes did in fact affect editorial policies; we suspect that a larger factor was editors' interest in increasing the citations to themselves and their editorial colleagues.

APPENDIX

Appendix: Listing of *American Economic Review* editors and editorial board members used for the citations-from-*AER*-commentary investigation conducted here and reported in the one-line table above.

***AER* list of names used for 1985**: Robert W. Clower, John G. Riley, George A. Akerlof, Clive D. Bull, Patricia Danzon, Michael R. Darby, Philip E. Graves, Jack Hirshleifer, Meir Kohn, Frederic S. Mishkin, Sherwin Rosen, Richard Schmalensee, Susan Woodward, Leslie Young.

***AER* list of names used for 1990**: Orley Ashenfelter, Robert H. Haveman, Bennett T. McCallum, Hal R. Varian, George Akerlof, James E. Anderson, Timothy F. Bresnahan, Henry S. Farber, Marjorie A. Flavin, Robert P. Flood, Claudia D. Goldin, Jo Anna Gray, Kenneth L. Judd, John H. Kagel, John F. Kennan, Dale T. Mortensen, Maurice Obstfeld, Edgar O. Olsen, John G. Riley, Richard Roll, David E. M. Sappington, Kenneth J. Singleton, Robert S. Smith, Barbara J. Spencer, Tresch Richard, John D. Wilson.

***AER* list of names used for 1995**: Orley Ashenfelter, Roger H. Gordon, R. Preston McAfee, Kenneth D. West, Alan J. Auerbach, Kyle W. Bagwell, David P. Baron, Rebecca M. Blank, Timoth F. Bresnahan, John Y. Campbell, H. Lorne Carmichael, Stephen G. Cecchetti, Lawrence J. Christianao, Nancy Gallini, Robert H. Haveman, Robert J. Hodrick, Kevin Hoover, R. Mark Isaac, Paul L. Joskow, John H. Kagel, John F. Kennan, John McMillan, Paul. R. Milgrom, Robert H. Porter, John Roberts, Paul Romer, Suzanne A. Scotchmer, Matthew D. Shapiro, Gary R. Solon, Jeremy Stein, Hal R. Varian, W. Kipp Viscusi, Carl E. Walsh, David W. Wilcox, Leslie Young.

***AER* list of names used for 2000**: Orley Ashenfelter, Tim J. Besley, R. Preston McAfee, Matthew D. Shapiro, James Andreoni, David P. Baron, Susanto Basu, Theodore C. Bergstrom, Francine D. Blau, Charles C. Brown, Allan Drazen, Dennis N. Epple, Timothy S. Fuerst, Jeffrey C. Fuhrer, Don Fullerton, Jordi Gali, Martin S. Gaynor, David Genesove, Gordon H. Hanson, Charles A. Holt, Adam B. Jaffe, Tracy R. Lewis, Robert A. Margo, Paul R. Milgrom, Robert A. Moffitt, Christina H. Paxson, Wolfgang Pesendorfer, Valerie A. Ramey, Michael R. Ransom, Sergio T. Rebelo, Jennifer F. Reinganum, Peter C. Reiss, Michael H. Riordan, Richard E.

Romano, David H. Romer, Andrew R. Schotter, Robert W. Staiger, Curtis R. Taylor, W. Kip Viscusi, Carl E. Walsh, David N. Weil, Kenneth D. West, Michael D. Woodford, Randall Wright.

REFERENCES

Brauninger, Michael and Justus Haucap. 2003. Reputation and Relevance of Economics Journals. *Kyklos* 56(2): 175-198.

Coelho, Philip R.P., Frederick De Worken-Eley III, and James E. McClure. 2005. Decline in Critical Commentary, 1963-2004. *Econ Journal Watch* 2(2): 355-361. Link.

Laband, David N. and Michael J. Piette. 1994a. The Relative Impacts of Economics Journals: 1970-1990. *Journal of Economic Literature* XXXII: 640-666.

_____. 1994b. Favoritism versus Search for Good Papers: Empirical Evidence Regarding the Behavior of Journal Editors. *Journal of Political Economy* 102(1): 194-203.

Laband, David N. and Robert C. Tollison. 2003. Dry Holes in Economic Research. *Kyklos* 56(2): 161-174.

_____. 2004. Dry Holes in Economic Research: Reply. *Kyklos* 57(4): 621-626.

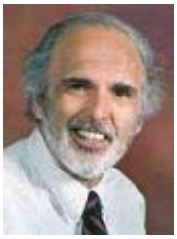
Laband, David N., Robert C. Tollison, and Gokhan Karahan. 2002. Quality Control in Economics. *Kyklos* 55: 315-324.

Mayer, Thomas. 2004. Dry Holes in Economic Research: Comment. *Kyklos* 57(4): 627-628.

Whaples, Robert. 2006. The Costs of Critical Commentary in Economics Journals. *Econ Journal Watch* 3(2): 275-282. Link.

Ziliak, Stephen and Deirdre McCloskey. 2004. Size Matters: The Standard Error of Regressions in the *American Economic Review*. *The Journal of Socio-Economics 33,* 527-546.

ABOUT THE AUTHORS



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

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

GO TO ORIGINAL ARTICLE BY COELHO ET AL (2005)

GO TO ARTICLE BY WHAPLES (2006)

Econ Journal Watch,
Volume 3, Number 2,
May 2006, pp 292-379.

DO ECONOMISTS REACH A CONCLUSION?

Do Economists Reach A Conclusion on Road Pricing? The Intellectual History of an Idea

ROBIN LINDSEY[*]

Abstract

ADAM SMITH WROTE SEVERAL PAGES ABOUT TOLL ROADS, AND Jules Dupuit developed a rich analysis of road pricing, but the English-language work by economists does not really get going until after Pigou (1920). The historical pattern of interest might reflect contemporary practice. During most of the nineteenth-century, toll roads were commonplace in Britain and the United States. They advanced social and economic goals. Although rail transport had displaced many of the trunk-line toll roads, what snuffed out the toll road was government policy. It is in the era of free-access highways, and especially after 1920 with the rise of the automobile and congestion, that economists had a growing problem to solve.

For decades after 1920, road pricing remained an ivory-tower idea. But in the 1990s interest grew significantly. Governments around the world became more supportive of road pricing. For example, the European Union has been promoting the application of marginal-cost pricing in

[*] Department of Economics, University of Alberta. Edmonton, T6G2H4.
In November 2004 I presented a preliminary version of this paper in a University of Alberta seminar and at the Southern Economic Association meetings in New Orleans. Helpful comments were received from both audiences. Thanks are also due to Ken Button, Don Shoup and Gabriel Roth for insightful comments. Any errors are wholly my responsibility. Financial support from the Social Sciences and Humanities Research Council of Canada (project "Road pricing in urban areas") is gratefully acknowledged.

transportation, and it has issued a number of policy papers[1] and funded a series of research projects. And in the United States the cause of road pricing has been advanced by the Value Pricing Pilot Program which funds innovative road and parking pricing measures for alleviating congestion.[2] The change in political attitudes towards road pricing may have been part of a broader ideological acceptance of market mechanisms. But another factor was technology: During the 1990s electronic toll collection was developed and implemented. The high transaction costs of old-fashioned toll collection and motorist delay were largely eliminated. Yet other reasons for the new interest in road pricing may have been the growing (or anticipated) revenue shortfalls from fuel taxes and other traditional sources, as well as recognition that other, indirect means of mitigating highway congestion were inadequate.

To get an idea of how the volume of research on road pricing has evolved over time, a search on a cluster of terms representing the idea of road pricing was performed using the ECONLIT database from its inception in 1969 through 2004.[3] Figure 1 shows a trickle of papers from 1969 to 1994 (annual mean 4.2), followed by a sharp increase to a much higher level for 1995-2004 (annual mean 21.4).[4]

The counts in Figure 1 do not fully reflect a number of recent books about road pricing, including studies of public and political acceptability (Schade and Schlag 2003, Ison 2004), case studies of the implementation process (Mylvaganam and Borins 2004, Richards 2005), an edited volume of contributions (Santos 2004) and mathematical/theoretical studies (Arnott, Rave and Schob 2005, Yang and Huang 2005), and ambitious scholarly

[1] See in particular European Commission (1995, 1998, 2001).

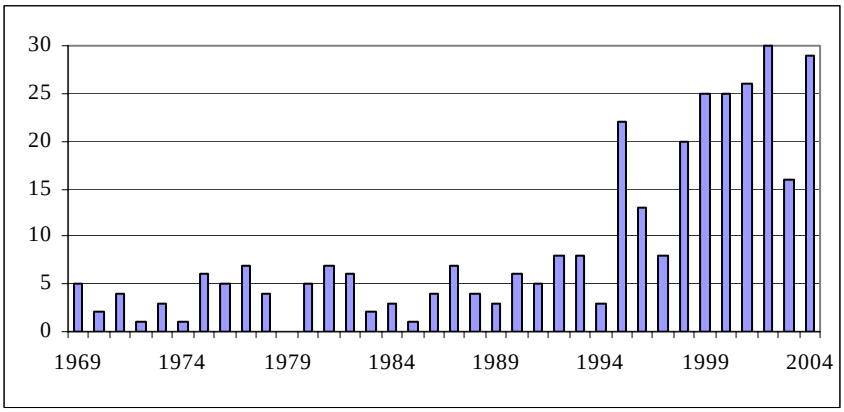
[2] See Value Pricing Pilot Program (2005), Transportation Research Board (2003) and DeCorla-Souza (2004).

[3] See Appendix A for details of the search procedure and coverage of the ECONLIT database. Year 2005 is excluded because entries for 2005 in the ECONLIT database were incomplete as of the end of January, 2006.

[4] Appendix B summarizes the coverage of road pricing in selected transportation economics textbooks since 1975. In contrast to Figure 1, the textbooks show no clear upward trend in the number of pages devoted to road pricing except for the incomplete draft of Small and Verhoef (2006). Button (2004) conducts a survey of research by all professions on road pricing and congestion in the United States and Europe. Over the period 1994-2002 he finds (Table 5.1) that of 953 doctoral dissertations on transportation in the US, only 50 relate to congestion. And for the category "social and behaviour studies on transport congestion in Europe" from 1996-2001 (Table 5.4), 60 percent of the studies dealt with measures to relieve congestion, but only 5.4 percent with road pricing. These statistics indicate that road pricing has not experienced any great upsurge of attention among transportation researchers as a whole over the last decade.

argumentation for toll roads (Roth 1996, Roth 2006).[5] The counts also do not include all of several recent special journal issues on road pricing.[6]

Figure 1: Publications on road pricing (1969-2004)



Source: ECONLIT database search on a cluster of terms representing the idea of road pricing

The flurry of interest in road pricing over the last decade or so might suggest that road pricing is an idea whose time has come (or, rather, come back!). Intercity toll roads are prevalent today in Western Europe, Mexico, Japan, China and other countries. And a number of projects are now either under development, or have been proposed, including area-wide tolling for Britain and the state of Oregon using Global Navigation Satellite Systems (GNSS) technology. Yet, despite the historical tradition of tolls and the longstanding theoretical argument for road pricing, only a few cities have

[5] In addition there is Shoup's (2005) authoritative and encyclopaedic book on parking which covers the technology and policies of parking pricing as well as the link between parking and road congestion.

[6] See "Recent studies on key issues in road pricing" (*Transport Policy*, 9(3), 2002), "Road pricing problems: Recent methodological advances" (*Networks and Spatial Economics*, 4(2), 2004), "Road pricing in practice and theory" (*Review of Network Economics*, 3(4), 2004), "The theory and practice of congestion charging" (*Transportation Research A*, 39A(7-9), 2005), "Road user charging: Theory and practice" (*Transport Policy*, 12(5), 2005), and "Research challenges in modelling urban road pricing" (*Transport Policy*, 13(2), 2006).

implemented road pricing in any form, and the number of schemes designed to control traffic congestion is even fewer.[7]

Although the idea wins intellectually, political acceptability remains a great challenge, and diverse attempts to introduce road pricing have failed politically.[8] The political dimension affects economists' judgments. Clarence Philbrook (1953) and W.H. Hutt (1971) favoured pursuit of the desirable regardless of political acceptability. But practical considerations lead many economists to focus on politically palatable reforms. The classic tension between the desirable and the politically acceptable is particularly relevant for highways, as we find ourselves in a freeway status quo that is difficult to undo.

The goal of this article is to assess whether economists, in their published judgments, agree that road pricing is a good idea.[9] An economist is defined to be someone with a postgraduate degree in economics or a job with a title of economist such as a teaching or research position at a university economics department.[10] Road pricing is defined broadly to include any form of direct user charges (e.g. tolls and area licenses), charges on urban and intercity roads, charges on any form of motorized transport, and charges for any purpose.[11]

---

[7] The major operational schemes (and their inception dates) are Singapore's electronic road pricing system (1998; a follow-on to the area-licensing scheme that began in 1975), toll rings in Norwegian cities (1986), London's congestion charge (2003), a handful of High-Occupancy Toll (HOT) lane projects in the US (1995), urban toll roads in Brisbane, Sydney and Melbourne (1999), and Highway 407 in Toronto (1999). On January 3, 2006, a toll cordon was launched in Stockholm. After a seven-month trial a referendum will be held, September 17, on whether to make the charge permanent.

[8] Notable failures include Hong Kong in the mid-1980s (Borins 1988), Rekening Rijden (Bill Riding) for the Randstad area in the Netherlands (Small and Gómez-Ibáñez 1998, §10.5.1), Cambridge UK (Oldridge 1995, Ison 2004), Edinburgh (McQuaid and Grieco 2005), several false starts in London prior to 2003 (Richards 2005), attempts during the 1970s to initiate congestion pricing demonstration projects in US cities (Elliott 1986, Higgins 1986), the Maine Turnpike (Colgan and Quinlin 1997), a section of the Trans-Canada Highway in New Brunswick (2000), and New York City (2002). Trondheim launched a toll ring in 1991, but the policy package in which the toll ring was embedded expired at the end of 2005, and the toll was terminated. Trondheim thus became the first city in modern times to stop collecting tolls.

[9] It is frequently claimed that they do agree; e.g. Small, Winston and Evans (1989, 86-87), Boyer (1998, 8), Thomson (1998, 99), Mohring (1999, 193).

[10] Articles in *The Economist* magazine are included although the authors are anonymous and may not have postgraduate economics degrees.

[11] See Appendix A for details on what is included as road pricing, and what is excluded. Various terms are used in the literature besides "road pricing," including "tolls," "road-use pricing," "road-user charging," "congestion charging," "congestion pricing" and "congestion

In short, economists do agree that highway congestion should be solved by pricing. Beyond that primary insight, however, there is much disagreement. Economists disagree over how to set tolls, how to cover common costs, what to do with any excess revenues, whether and how "losers" from tolling previously free roads should be compensated, and whether to privatize highways. These disagreements fill a lot of pages, while the main point of agreement is largely taken for granted.

It is not easy to assess agreement on the fine points. An economist may approve of road pricing in principle, but balk at particular schemes because of high administration costs, inequity, or other reasons. Another difficulty is that many authors never voice a judgment about road pricing, but develop results that seem to point toward a judgment. Another problem is that there is a bounty of material, and anything short of a book must be selective. Yet another frustration is that major analytic and practical advances have been made by scholars from elsewhere, notably civil engineering. Also, there have been exchanges between economists and others that have influenced the progression of thinking. Here, some of the work of non-economists will be mentioned while taking care to identify them as such.

There have been a number of surveys of attitudes towards road pricing. Most of these are opinion surveys of the public and/or policy-makers.[12] I am aware of two recent surveys of economists: one by Ison (2004, Ch. 3), an economist, and one by Gulipalli, Kalmanje and Kockelman (2005), who are engineers. The literature survey here differs in covering a larger number of economists and in adopting a historical approach. It also assesses opinions on the basis of published work, whereas Gulipalli et al. (2005) use written attitudinal questionnaires.[13]

---

metering." I use "road pricing" to encompass all these terms. Although freight transport charges and the land-use effects of road pricing were included in the survey, they do not appear in the paper. In part this is because understanding of these important topics is still immature, and (in the case of freight transport) because relatively little about it has been written by economists.

[12] See for example Harrington, Krupnick and Alberini (2001), Berg (2003), Johansson et al. (2003), Ison (2004) and Farrell and Saleh (2005).

[13] Three results of the Gulipalli et al. (2005) survey will be mentioned later. Although neither Ison (2004) nor Gulipalli et al. (2005) offer systematic comparisons between economists and other people, their results are consistent with received wisdom that economists are more favourably disposed towards the use of economic instruments such as road pricing. This was confirmed by Brittan (1973) in a survey of attitudes towards peak-period public transport fares—a close cousin of congestion pricing. Baumol and Fischer (1987, 383) summarize Brittan's findings: "Some years ago, Samuel Brittan (1973, p.93) conducted a survey on a

INTELLECTUAL EVOLUTION

Other scholars have reviewed the theory of road pricing.[14] My review treats Adam Smith and Jules Dupuit, and then dives into the twentieth century and continues to recent developments.

Adam Smith

Adam Smith devoted several pages of *The Wealth of Nations* to transportation projects, notably "high roads," which in his day were operated individually by a "trust," a local, independent authority voluntarily financed by bonds. Smith clearly favoured that all such "publick works" "be so managed, as to afford a particular revenue for defraying their own expence, without bringing any burden upon the general revenue of the society" (1937, 682). The first argument made by Smith is one of equity: "When the carriages which pass over a highway or a bridge ... pay toll in proportion to their weight … they pay for the maintenance of those public works exactly in proportion to the wear and tear which they occasion of them. It seems scarce possible to invent a more equitable way of maintaining such works" (683). But Smith also makes several practical arguments. First, he makes a kind of knowledge and accountability argument in matching supply to demand: "When high roads, bridges, canals, &c. are in this manner made and supported by the commerce which is carried on by means of them, they can be made only where that commerce requires them, and consequently where it is proper to make them" (683). Smith elaborates on how financial independence is a guard against extravagance and misplacement (687). He examines the proposal to

variety of policy issues, questioning 102 economists and 91 Members of parliament in Great Britain. When asked 'in order to make the most efficient use of a city's resources, how should subway and bus fares vary during the day?' predictably, 88 per cent of the economists agreed that 'they should be relatively high during rush hour to minimize the amount of equipment needed to transport the daily traveller'. But 60 per cent of the Conservative MPs and 39 per cent of the Labour MPs chose the answer that "they should be the same at all times' and 40 per cent of the Labourites felt 'they should be relatively low during [the] rush hour'. This is one of many examples where policies recommended enthusiastically by economists on grounds of efficiency are not quite as welcome to other people, who often question them because of their apparent inequity."

[14] See in particular Hau (1992) reprinted as Hau (2005a,b), Roth (1996, Annex to Chap. 4), Thomson (1998), Ison (2004), Richards (2005), and Toll Roads News.

have the central government take over operation of the turnpikes. He opposes the proposal on the grounds that the state would grow dependent on the toll revenues and increase the tolls unduly, encumbering commerce and driving up prices to the final consumer, and that the government is more likely to neglect maintenance. Smith says that it will sometimes be the case that a project is socially desirable and yet unable to finance itself entirely by tolls. In such cases he again presses the principle of local authority and financing, for virtues in accountability, local knowledge and consideration, and correction of errors (689). Throughout, Smith reminds us that if a work does not support itself, then it must be supported by other means that pose "a very considerable burden" (768).

Smith's discussion represents a comparative-institutions approach. The main principle in his discussion seems to be the independence and autonomy of the facility. Consistently applied, the principle would seem to point to full propertization of the facility. Smith describes (684) how a canal in France was turned over to the project's engineer, apparently essentially as private property, and how his residual claim and authority induced good maintenance and good results, better than if the canal had been in the hands of commissioners that would have no such interest. In turning to the case of high roads, however, Smith changes his tune. Unlike a canal, a road wanting in maintenance does not become altogether impassable. "It is proper, therefore, that the tolls for the maintenance of such a work should be put under management of commissioners or trustees" (684). Smith's reasoning here is odd. He seems to believe that private ownership will deliver adequate maintenance if doing nothing would make the facility unusable, but not if skimping on maintenance merely degrades the quality of service.[15] Furthermore, private ownership does not preclude accountability to local public officials concerning maintenance.

Jules Dupuit

Dupuit was one of a group of French engineers who wrote extensively during the nineteenth century about transportation and other areas of economics. Ekelund and Hébert (1999, 3) sum up their contributions: "It is fairly easy to make the case that the state engineers of

[15] This reasoning is at odds with that of Knight (1924) and later writers mentioned below.

the Corps des Ingénieurs des Ponts et Chaussées pioneered the field of transport economics."[16]

Dupuit understood most of the ideas underlying Pigouvian tolls and other aspects of road pricing, including consumer's surplus and the importance of product quality generally—as well as travel time specifically. He built formal models and derived the toll that would recover costs and the profit-maximizing toll for a monopoly, and analyzed at length the potential for price discrimination to boost revenues. He identified the danger that tolling one (congestion-free) bridge would divert traffic onto an alternative less attractive bridge (Dupuit 1952 [1844], 105). Dupuit also showed an appreciation for the wider public finance implications of tolls: "if I had wished to treat exhaustively of only one question, namely, whether or not to establish tolls, I would have had to examine by what new tax or what increase in taxation tolls could be replaced and what would be the effects of these taxes; I would have been led into a full-fledged theory of taxation" (1952 [1844], 30-31).

Dupuit did not appreciate the distinction between short-run and long-run marginal cost, and did not address congestion (which presumably was a much smaller problem given the vehicles of the day). He saw tolls more as a means of covering long-run costs than of managing efficient usage: "If… the bridge is public property, the government will want to recover from the toll merely a fixed sum representing interest on the capital spent for construction, maintenance cost and perhaps amortization" (1962 [1849], 11).[17]

Thus, we might interpret Dupuit as being in the company of later economists who favoured average-cost pricing. And, given Dupuit's support of markets and the principle of user pays, he appears to be an early sympathizer—if not a proponent—of private toll roads.

---

[16] Ekelund and Hébert (1999) refer to Dupuit and others of the school as "engineers," "econo-engineers" and "economists". Derycke (1998) also refers to "economist-engineers". Thus, it seems reasonable to classify Dupuit as an economist for the purpose of this survey.

[17] As quoted by Ekelund and Hébert (1999, 180). Derycke (1998, 63) draws a similar conclusion: "While the first school, which dates back to the mid-19th century but has much earlier roots, has developed what is more a theory of *funding tolls*, the second, founded sixty years later in Britain, has focused on *decongestion tolls*."

Twentieth century leaders

Arthur Pigou

It appears that Pigou deserves credit for suggesting tolls on public roads to alleviate congestion. Pigou's (1920) famous example of two parallel roads occupies just one paragraph:

> Suppose there are two roads ABD and ACD both leading from A to D. If left to itself, traffic would be so distributed that the trouble involved in driving a 'representative' cart along each of the two roads would be equal. But, in some circumstances, it would be possible, by shifting a few carts from route B to route C, greatly to lessen the trouble of driving those [sic] still left on B, while only slightly increasing the trouble of driving along C. In these circumstances a rightly chosen measure of differential taxation against road B would create an 'artificial' situation superior to the 'natural' one. (Pigou 1920, 194) [18]

Frank Knight and others

Knight (1924) criticized Pigou on the grounds that an efficient outcome in his two-roads example would result without a tax if roads were privately owned. Knight built on Pigou's example by assuming that route B is smooth—but narrow and congestible, while route C is rough—but wide enough to accommodate all potential traffic without delays. Knight then observed:

> If the roads are assumed to be subject to private appropriation and exploitation, precisely the ideal situation which would be established by the imaginary tax will be brought about through the operation of ordinary economic motives….the owner of the narrow road can charge for its use a toll representing its "superiority" over

[18] The two-roads example is sometimes incorrectly attributed to Pigou (1912); e.g. Buchanan (1956, 163), Derycke (1998), Mills (1981, fn 1), Newbery (1988 and 1989).

> the free road, in accordance with the theory of rent, which is as old as Ricardian economics. (Knight 1924, 586-587)

As Buchanan (1956), Edelson (1971) and Mills (1981) showed decades later, Knight is correct only under stringent assumptions. Mills provides a helpful summary:

> It has long been known that Knight's claim for the generality of this result was excessive. James Buchanan showed that it held in the road example only because there was an alternative route to the congestion-prone road, and that ownership rights could be conferred without simultaneously, and inadvertently, conferring monopoly power. After describing some alternative institutional arrangements where the result fails to hold, he concluded that private ownership can be relied upon to achieve efficient resource use '[o]nly in those cases where the extent of commonality of usage is limited to a relatively small proportion of the total resource supply…' (p.315). That is, there must be a sufficient supply of alternatives to the facility in question to prevent monopoly power from existing if it is owned privately. Later, working with a variation of the road example, Noel Edelson qualified Knight's claim in another way. He showed that it depended on all users of the two roads having the same imputed value of time spent in transit. Otherwise, over- or undercongestion would result. (Mills 1981, 493)

Little further was written about road pricing until the mid-1950s. Clark (1923, 304) advocated user charges in response to problems of highway damage, unfair competition vis à vis railroads, and the need to pay for highways. Peterson (1932) maintained that roads should be priced like other commodities. Buchanan (1952) covered many of the pros and cons of tolling. He emphasizes the potential role of tolls in improving usage: "[C]oncentration on the allocation of benefits and thus the equitable means of distributing total highway costs has all but obscured the far more important problem of adjusting user charges in order to promote an optimum utilization of an existing highway system" (Buchanan 1952, 98).

Hold it! Ronald Coase's comparative-institutions challenge

In his 1946 article "The Marginal Cost Controversy," Ronald Coase took exception to marginal-cost conclusions arising from the model-based work of Harold Hotelling, Abba Lerner, James Meade, and J.M. Flemming. This literature does not treat highways in particular, but the general discussion applies to highways.

Model-based discourse boils reality down to a model, solves for optimality, and then tends to emphasize the implications of that optimality exercise. Coase's approach, also characteristic of Adam Smith, is that of *comparative institutions*, whereby scholars formulate and compare institutional alternatives one against another, in light of whatever considerations seem relevant and important, including especially the limitations and likely failings of the governing institutions themselves (see also Demsetz 1969). The comparative assessment appeals to broad, never-fully-identified sensibilities, rather than to a formal model that presumes to capture all that matters to coming to a judgment.

The thrust of Coase's article is to suggest that there are knowledge and accountability advantages to organizing the facility in such a way that it must support itself. That assures that only projects likely to pay for themselves will be undertaken. Coase emphasizes that such arrangements provide a substitute for the need of masterful knowledge, and cites Hayek on socialism (170). Such arrangements focus the mind and enhance the planning accuracy of those accountable. The alternative "marginal-cost"/subsidization arrangement involves not only distortions and troubles in taxation of incomes etc. (178-9), but does not provide, even after the fact, a basis for judging the worthiness of the facility (176). As Smith noted, subsidized facilities will be subject to political factors and overproduction. Coase recognizes that self-financing might result in some facilities not being built that should have been (181). His preferred form of self-financing—multi-part pricing[19]—would however provide a more refined method of appropriating value than simple average-cost pricing, and therefore would reduce the bias towards underproduction. Coase feels that the self-financing principle is so important, however, that even if it came down to a choice between marginal-cost pricing (and subsidies) and average-cost pricing (no subsidies), there should be no presumption in favour of marginal-cost pricing.

[19] Coase notes that this solution was suggested earlier by C.L. Paine (1937) and by E.W. Clemens (1941).

In Coase's article, the question of ownership is somewhat vague. He seems to say that if the facility is a government enterprise, then the facility ought to have to pay for itself, and its managers ought to look to multi-part pricing. Coase is silent on whether he would prefer private ownership, and on whether privately-owned facilities ought to be regulated by the government. His other works, however, certainly seem to favour private ownership.

Mid-twentieth century models and practical proposals

During the late 1950s and early 1960s, the cause of road pricing was advanced by models and practical proposals, with economists at the vanguard. Beesley and Roth (1962-63) described the emergence of a new role for economists: "Growing traffic congestion and recent advances in computing techniques have induced economists to look at the problem in detail, and thus to move into a field that has hitherto been the preserve of engineers and town planners" (Beesley and Roth 1962-63, 184).

Looking back at the period Thomson (1998, 94) writes:

> Until the 1960's roads were regarded as the province of engineers. The planning of urban roads and road traffic was carried out by civil engineers, with occasional assistance from architects. The professions of urban planning and transport planning were in their infancy. Engineers were in control and saw no need for economists. The intrusion of economists into teams of transport planners was often treated with a mixture of suspicion and amusement. (Thomson 1998, 94)

William Vickrey

Drèze (1994) and Arnott (1998) provide incisive appraisals of Vickrey's extensive writings on marginal-cost pricing and traffic congestion. Between 1948 and his death in 1996, Vickrey wrote some 40 articles that treated most road-pricing issues. His 1948 paper conveyed two ideas that ran through all his later work. One was that price should be set at short-run marginal cost (SRMC) rather than either long-run marginal cost (LRMC) or

average cost (AC).[20] The second was that random demand fluctuations should be met with responsive pricing whereby prices are adjusted to match SRMC as closely as practical. Vickrey was also interested in the technology for road pricing, and in 1959 he proposed a system of network-wide tolls for Washington, DC. His interest was spurred by the great waste that he saw in pricing of road transport:

> I will begin with the proposition that in no other major area are pricing practices so irrational, so out of date, and so conducive to waste as in urban transportation. Two aspects are particularly deficient: the absence of adequate peak-off [sic] differentials and the gross underpricing of some modes relative to others. In nearly all other operations characterized by peak-load problems, at least some attempt is made to differentiate between the rates charged for peak and for off-peak service. Where competition exists, this pattern is enforced by competition: resort hotels have off-season rates; theaters charge more on weekends and less for matinees. Telephone calls are cheaper at night. . . . But in transportation, such differentiation as exists is usually perverse. (Vickrey 1963, 452)

Alan Walters

Like Vickrey, Walters supported SRMC pricing, but in his early writings (Walters 1954) he focused more on the fallacy of AC pricing and misguided investment. Walters' most widely cited work is his 1961 article on congestion and congestion pricing. Perhaps surprisingly, his support there for road pricing is rather qualified (683-5), and at one point (685) he remarks "Fuel taxes are probably the most useful form of deterrent."

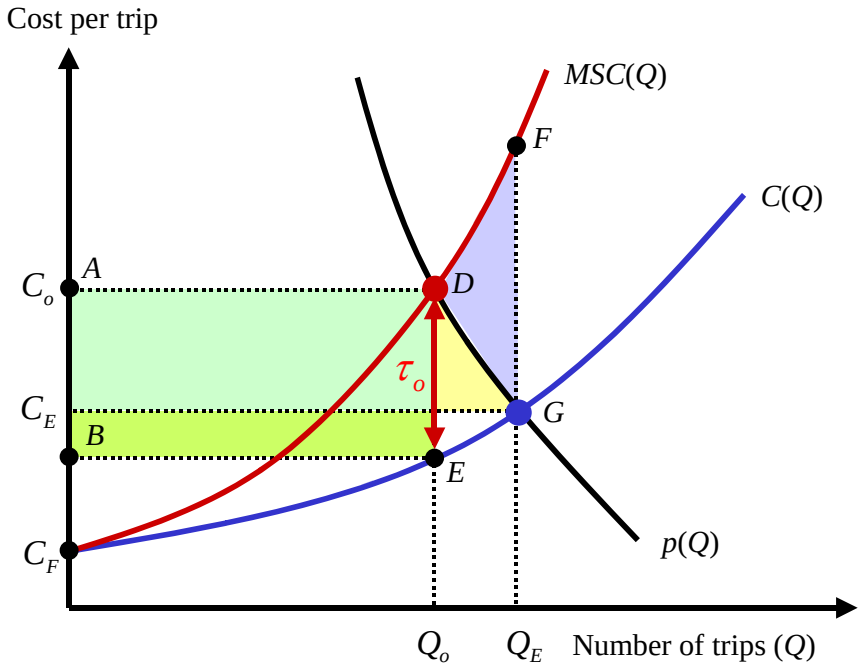
These misgivings notwithstanding, Walters (1961) serves as a useful landmark. His approach, which will be referred to here as the *basic model*, can be illustrated with what is called the *conventional diagram*, shown in Figure 2.[21] Individuals are assumed each to make one trip, one person per vehicle, along a single stretch of road between a common origin and a common

[20] A good summary of his arguments is found in Vickrey (1987).

[21] The presentation here follows Lindsey and Verhoef (2001). Figure 2 is a highly modified version of Walters (1961, Figure 1).

destination. The number of trips, measured as an hourly flow, is plotted on the horizontal axis. The cost per trip, which includes vehicle operating costs and the opportunity cost of travel time, is plotted on the vertical axis. As the number of trips increases, congestion forces drivers to slow down and the average cost of a trip, $C(Q)$, rises. Because each motorist accounts for a negligible fraction of flow and all trips are identical in cost, the motorist's marginal cost coincides with average cost. The function $C(Q)$, then, represents both the average cost and the private marginal cost.

Figure 2: The basic road pricing model



Demand for trips is described by a conventional downward-sloping inverse demand curve, $p(Q)$. Without a toll, equilibrium occurs at the point of intersection, G. Q_E trips are taken each at a cost C_E. The equilibrium is inefficient because individuals disregard the delay they impose on other

travelers. The total social cost of Q trips is $TC(Q) = C(Q) \cdot Q$, and the marginal social cost of a trip is $MSC(Q) = \partial TC(Q)/\partial Q = C(Q) + \partial C(Q)/\partial Q \cdot Q$. The social optimal is found at the point of intersection, *D*, of $MSC(Q)$ and $p(Q)$. The optimal number of trips, Q_o, is less than Q_E. To support the optimum, travellers must be made to face a cost of C_o. This can be accomplished by imposing a toll of $\tau_o = MSC(Q_o) - C(Q_o) = \partial C(Q_o)/\partial Q_o \cdot Q_o$ equal to the marginal external congestion cost of a trip. The welfare gain from imposing the toll as measured by the increase in social surplus is given by area *DFG*.

The basic model is attractive in several respects. It is straightforward and amenable to graphical description. The welfare gain is represented visually. The toll is familiar to contemporary readers as a Pigouvian toll. And the formula for the optimal toll is intuitive since it equals the marginal delay imposed by a driver on each other driver, $\partial C(Q_o)/\partial Q_o$, times the number of other drivers, Q_o. These aspects of the basic model have no doubt helped to sell economists on road pricing. But further inspection of Figure 2 reveals four sobering facts:

> (1) Tolling raises drivers' private costs, as indeed it must if travel is curtailed. The revenue from the toll accrues to the toll-road operator, which is usually assumed to be a government agency. Unless the government uses the revenue to expand road capacity, to improve an alternative form of transport, to reduce other user charges, or to provide rebates to drivers in some lump-sum fashion, drivers end up worse off.
>
> (2) Toll revenues are just a transfer from travelers to the operator, but toll collection entails infrastructure, operating and administration costs as well as inconvenience for travelers. Furthermore, as Figure 2 is drawn, the toll revenue, *ADEB*, is large compared to the welfare gain, *DFG*, which is likely to be the case if travel demand is price inelastic as it typically is at peak times. Hence, unless the costs of toll

collection per dollar of revenue garnered are small, the net social benefit from tolling will be negative. [22]

(3) Congestion is not eliminated because the cost of travel at point *E* exceeds the cost under free-flow conditions, C_F. Consequently, tolling can only be marketed to the public as a way of reducing congestion, not eliminating it.

(4) The optimum at point *D* cannot be deduced just by observing the no-toll equilibrium at point *G*. To solve mathematically for *D* and the supporting toll τ_0, it is necessary to know, or to estimate, the *C*(*Q*) and *p*(*Q*) curves.

The basic model is very simple, including as it does a number of simplifying assumptions:

- a single road connecting one origin to one destination;
- one individual per vehicle;
- vehicles contribute equally to congestion;
- individuals are identical except for their reservation price to make a trip;
- traffic flow, speed and density are uniform along the road, and are independent of time;
- congestion is the only market failure; i.e. there are no other transport externalities or distortions elsewhere in the economy; and
- there are no shocks due to accidents, bad weather, special events, etc.

It is not surprising that these simplifying assumptions, combined with the discouraging properties of the equilibrium, left the basic model open to criticism from engineers as well as some economists.

[22] However, most other sources of public revenue are also costly to collect, and unlike congestion tolls they create deadweight losses. From this wider perspective congestion pricing may be beneficial even if the collection costs are significant.

An aside on hypercongestion

Figure 2 omits one element of Walters' basic model, namely, the backward-bending or hypercongested branch of the average cost curve, $C(Q)$. This omitted branch corresponds to travel conditions, frequently observed on highways, in which travel speed and flow (treated as number of trips in Figure 2) are positively correlated—and travel time (and cost) and flow are correspondingly negatively correlated. In other words, congestion is so heavy that flow declines. The speed at which flow peaks depends on lane width, geometrics, spacing between on-ramps and off-ramps and other highway characteristics. Because of considerable scatter in speed-flow data it is often difficult to identify a precise value.[23] The circumstances under which hypercongestion occurs are still not fully understood although reductions in capacity downstream of the point of observation and turbulence created by vehicles entering from on-ramps and exiting from off-ramps have been identified as contributory factors.

Hypercongestion cannot be captured by a static model of the sort depicted in Figure 2. Small and Chu (2003, 342) explain the logic: "Hypercongestion is a real phenomenon, potentially creating inefficiencies and imposing considerable costs. However, it cannot be understood within a steady-state analysis because it does not in practice persist as a steady state. Rather, hypercongestion occurs as a result of transient demand surges and can be fully analysed only within a dynamic model. Even if the dynamic model is converted to a static one through the use of time averaging, the appropriate specification of average cost depends on the underlying dynamics. In virtually all circumstances that specification will portray average cost as a rising function of quantity demanded, even when hypercongestion occurs."[24]

---

[23] Various empirical studies and speed-flow models are reviewed in Small and Verhoef (2006, Chapter 3).

[24] These arguments notwithstanding, it does appear that tolling may be particularly beneficial at times when hypercongestion would otherwise occur. (Simulation results by several authors using different dynamic models are reviewed in Small and Verhoef (2006, Chapter 4).) Moreover, traffic count data on State Route 91 in California show that the High Occupancy Toll (HOT) lanes support higher hourly traffic flows per lane than do the toll-free general-purpose lanes (Poole and Orski 2003, 6). However, this behaviour has yet to be systematically analyzed and explained, and the idea that congestion pricing can increase throughput on roads generally is controversial. One consideration is that, unlike the HOT lanes, the toll-free lanes have intermediate entrances and exits between their end points that could create turbulence in the traffic stream.

Martin Beckmann, Bartlett McGuire and Christopher Winsten

Having presented the conventional model of Walters (1961), I now step back in time to a more comprehensive work of 1956. In a pathbreaking monograph that covered rail as well as road transport, Beckmann, McGuire and Winsten (1956) showed how mathematical programming methods could be used to solve for traffic equilibrium on a road network and how to use tolls to manage route choices and an optimal number of trips.[25] In a sense, their work extends the insights of Walters (1961) to a network. Like Vickrey and Walters, Beckmann et al. were interested in tolls to manage efficient usage of roads rather than to finance them:

> In this discussion, tolls are looked upon, not as a means of financing road construction, but as a means of bringing about the best utilization of the highway network. This is in keeping with the growing acceptance among modern economists of the proposition that best use of facilities requires methods of pricing the services of these facilities that reflect the incremental cost attributable to each service demanded by an individual user. Because of the non-linearity in the relation between amount of use and cost, such pricing does not necessarily produce revenues equal to the total cost of operating and financing the facility. (Beckmann et al. 1956, 8 [Introduction by T.C. Koopmans])

Herbert Mohring

A theme that underlies Mohring's extensive writings on transport economics is that economic principles of price and value are applicable to transportation generally, and highway travel specifically.[26] Like other leaders of the era, Mohring was a strong advocate of SRMC pricing. His most celebrated result, derived in Mohring and Harwitz (1962), is the *cost-recovery theorem* that the revenues from SRMC pricing just suffice to pay for optimal capacity if capacity is perfectly divisible and supplied at constant marginal

[25] Boyce, Mahmassani and Nagurney (2005) provide an insightful review of the impacts of Beckmann, McGuire and Winsten.
[26] See, for example, Mohring (1976, Ch. 1).

cost, and user costs are homogeneous of degree zero in usage and capacity. The cost-recovery theorem is especially noteworthy for establishing that, despite the misgivings of Beckmann et al. noted just above, there may be no conflict between SRMC pricing and AC pricing.[27] Put another way, there may be no conflict between Dupuit and Pigou.

Anthony Downs

Downs' writings on traffic congestion span more than 40 years. His attitudes towards road pricing will be discussed later. In Downs (1962) he describes why attempts to alleviate congestion by building new roads or expanding existing ones are likely to be frustrated as travellers shift from other routes, other times of day and other modes until the new capacity is clogged. Downs formalized this behaviour as: "Downs's Law of Peak-Hour Traffic Congestion: *On urban commuter expressways, peak-hour traffic congestion rises to meet maximum capacity*" (Downs 1962, 393).

Although Downs did not analyze tolls, his article helped to motivate road pricing because Downs's Law seemed to apply to all policies that improve traffic conditions except policies that operate via the price mechanism.

The Smeed Report

In 1962 the UK Ministry of Transport assembled a panel of experts and asked them to examine different forms of road taxes.[28] The result was the Smeed Report (UK Ministry of Transport, 1964).[29] Like Vickrey, the

[27] In the same year, Nelson (1962, 435 [italics added]) had complained that "the loosening of all constraints of user revenue coverage of total costs on wasteful investment under the marginal-cost pricing scheme would work against a more efficient resource allocation; and *it has yet to be shown* that sufficient investment in economically-justifiable public transport facilities could be assured without the revenue-coverage criterion." As Mohring (1964, 1) observed after citing Nelson, "This and similar charges that fiscal irresponsibility is involved in advocating congestion pricing for highway services have been strongly refuted."

[28] The authors considered it outside their terms of reference to discuss the disposition of surplus revenues from road pricing. I am grateful to Gabriel Roth for making this point.

[29] The Smeed Report is included in this survey because many of the panel members were economists who reportedly thought along similar lines. Braybrooke (1974), a political scientist, describes the Smeed Committee members as advocating the same proposal (p.24): "I shall treat Beesley and his fellow-economists as one station, in spite of their not being

Smeed Report emphasized the practical aspects of road pricing, although it put more emphasis on user-friendliness. The conclusions of the report were upbeat:

> When we started our work, we set out a list of 17 requirements which we considered desirable for a road pricing system. Some at least of the six meter systems show promise of satisfying all these requirements. The main conclusion that emerges from our work, therefore, is that there is every possibility that at least one of these proposals could be developed into an efficient charging system and could yield substantial benefits on congested roads. (UK Ministry of Transport 1964, 42) [30]

Gabriel Roth

Roth was trained as a civil engineer and as an economist, and served on the Smeed Committee panel. Several themes run through all his major works (e.g. Roth, 1966, 1967, 1996):

- economic principles should be applied to roads;
- road pricing should be used both as a Pigouvian tax and as a means of funding road construction and maintenance;
- roads should be controlled by a roads authority that behaves like a competitive market; and
- problems should be approached as an issue of comparative institutions, not solving a model.

Roth did not endorse private roads in his early writings, but he later became more favourable towards them (e.g. Roth 2006).

organized as a group, since they were throughout the period in close touch with one another, supporting each other's arguments and advocating … the same proposal."

[30] The Smeed Report was not universally welcomed. Goodwin (1997, 2) remarks how "a now retired civil servant told me that as a young man he found in the Ministry of Transport files a note, in the personal hand of the Prime Minister, Sir Alec Douglas Home, saying 'let us take a vow that if we are re-elected we will never again set up a study like this one'."

Late twentieth-century extensions

The insights into congestion pricing were extended in various directions over the next thirty years or so. One important extension was to develop dynamic models that treat in a conceptually satisfactory manner temporal peaking of travel demand and the transitory nature of congestion that can be manifest as hypercongestion. Another stream of research explored the robustness of the cost-recovery theorem to relaxation of assumptions.[31] An especially significant contribution was to add road damage to the basic model, and to derive rules for efficient pricing of pavement damage and optimal pavement durability in combination with rules for efficient pricing of congestion and optimal road capacity. The key pieces are Newbery (1988 and 1989), Small and Winston (1988), and Small, Winston and Evans (1989). Roughly speaking they established that, under plausible assumptions, the sum of congestion and road damage charges just pay for the combined costs of optimal highway capacity and maintenance.

A question of practical interest that arose was whether a social optimum can be supported by tolling road links, or whether tolls need to be differentiated according to the origins and destinations of travellers. Dafermos and Sparrow (1971), two specialists in operations research, claimed that origin-destination-based tolls are required. As Littlechild (1973), an economist, pointed out (195) this would "render totally impracticable a road pricing system". But Littlechild showed that the Dafermos-Sparrow result holds only if travel times are equalized on all feasible travel paths, which is unlikely to be the case in practice.

Yet another stream of research examined under what conditions efficient pricing can be implemented with anonymous link tolls; i.e. tolls that do not depend on the characteristics of vehicles, motorists or trips. A key result, derived by Arnott and Kraus (1998) using a dynamic model, is that anonymous link tolling is efficient if tolls can be varied freely over time. These and other developments strengthened the modeling foundations of SRMC pricing as well as the case for launching road congestion pricing experiments.

Reluctance to do the obvious

The research on road pricing in the early 1960s was taken seriously enough by the US Highway Research Board that it organized a panel

[31] See Hau (1998 and 2005b) for a review.

session on congestion pricing and several papers were published in the 1964 *Highway Research Record*. Two of the papers were distinctly hostile. Grubbs (1964), an economist, targeted his critique at Walters (1961) and concluded that "assumptions underlying the proposition are too improbable to serve as a foundation for public policy … for highways in the United States" (Grubbs 1964, 15).

St. Clair (1964), a highway engineer, asked "three economists of the Bureau of Public Roads" (82) to comment on Walters' article.[32] The economists each raised a number of concerns. Selected and representative passages from their assessments follow:

> Professor Walters on page 677 assumes that traffic is homogeneous, all drivers are the same, and all vehicles have the same costs and speed, etc. He agrees that this is wholly unrealistic, but pursues his model in terms of these assumptions as though these were not people with individual preferences but water moving through a conduit pipe. (*Sidney Goldstein*, quoted by St. Clair 1964, 83)

> Presumably the purpose underlying the toll suggestions is to maintain and increase concentration of activity at the urban core. I doubt if a more self-defeating proposal—in the long run—could be devised. (*E.L. Kanwit*, quoted by St. Clair 1964, 84)

> [T]here is no way possible to even make a reasonable estimate of the necessary data. Hence, no practical purpose is served—neither is any theoretical validity obtained. (*John Rapp* [regarding the feasibility of using Walters' equations], quoted by St. Clair 1964, 85)

In the same issue of *Highway Research Record*, Zettel and Carll (1964)—two engineers—provide a more balanced assessment of road pricing. Nevertheless, after several pages of analyzing the various possible impacts of tolls, they remark:

[32] St. Clair does not say whether the economists had postgraduate degrees. But it seems reasonable to suppose that they had a job title of "economist" and therefore meet the definition of an economist adopted for this survey.

> At this point, one suffers mental indigestion trying to picture the tolled, the tolled-off, and the untolled, the users and the nonusers, bouncing around among the alternatives, all the while a blinking giant of a computer is fixing and refixing tolls, shadowing users, and redistributing income to promote the general welfare through an optimal arrangement, not only of travel, but also of nontravel. (Zettel and Carll 1964, 60)

Turvey (1968, 101) voiced similar criticism of the simplifying assumptions used in the standard peak-load pricing model: "[T]hey remove from the discussion some of the most interesting and important issues—in the field of electricity supply, at any rate." And "[W]hat follows is an attempt to bring [the authors] down from their ivory towers." In his concluding paragraph, Turvey remarks (113): "The theoretical 'solutions' to the peak-load problem are a beginning, not an end, serving to dispose of past confusion about the principles of allocating cost. While the matters which then have to be examined are less suited to the tools of the armchair economist, they are both important and fascinating. The only practicable method of securing quantified analysis is to engage in large-scale long-term experiments of the type now being pioneered in Britain." Turvey's remarks reveal some clear parallels between electricity and road pricing.

These and other criticisms were mostly focused on the practical relevance of the model. As already noted, the modeling literature has been elaborated considerably since the 1960s. Technology has advanced greatly too. But implementation of road pricing in urban areas is still quite limited. Economists have been perplexed by this failure:

> Seldom has applied economics produced an idea with such unanimous professional conviction in both its validity and its political unacceptability. A.A. Walters's article on 'congestion' in the *New Palgrave Dictionary of Economics* states flatly, 'The best policy to deal with urban road congestion is likely to be some form of road pricing. However, road pricing is the exception rather than the rule.' The consensus among professional economists in favour of this approach on economic grounds is strong. The theory is now refined and standard; implementation has been widely explored; numerous empirical studies have predicted its effects; and the whole package has made its

> way into standard textbooks in urban and transportation economics. (Small, Winston and Evans 1989, 86-87)
>
> It has been a commonplace event for transportation economists to put the conventional diagram on the board, note the self-evident optimality of pricing solutions, and then sit down waiting for the world to adopt this obviously correct solution. Well, we have been waiting for 70 years now, and it's worth asking what are the facets of the problem that we have been missing. Why is the world reluctant to do the obvious? (Lave 1995, 465)

Maybe part of the problem has been the emphasis on models. When the intellectual focus is on capturing the universe in a model and solving for optimality, the intellectuals tend to neglect many practical human and institutional arguments that ought to weigh heavily in a judgment between alternative arrangements. The model-based road-pricing literature has neglected the points of Smith and Coase about the organizational, accountability, and knowledge virtues of an independent facility that must recoup its expenses itself. Coase's sensible suggestion of multi-part pricing has not received much attention, probably because it is not so amenable to graphical representation or simple mathematical formulation. Since the 1960s, model-based discourse has enjoyed great prestige, leaving others such as Gabriel Roth less heeded and less influential.

ATTITUDES TOWARDS ROAD PRICING ON PUBLIC ROADS

Economists have exhibited diverse attitudes about the various aspects of road pricing. This section is organized by such aspects.

Treating roads like other market goods

One argument for road pricing is simply that roads should be treated like other goods, and road users impose costs on other users as well as on society as a whole. As noted above, this argument underlies Mohring's

work. It is also apparent in the following quotes, arranged in chronological order, including passages written by Murray Rothbard and Walter Block—strong proponents of private roads.

> While there are those who maintain that transportation is so peculiar an industry that it is not properly amenable to the principles underlying the price mechanism, the majority of economists probably accept the view that, in general, a transport facility to be justified must be able to pay its way. (Peterson 1932, 425) [33]

> This book attempts quite simply to apply to the commodity 'road space' the economic principles on which we rely for the allocation of most of our goods and services. Its objects are to show that traffic congestion can be sensibly dealt with only if the economic factors that underlie it are understood, and to stress a number of points, which, though not new, are not generally recognized. (Roth 1967, 11)

> Frantically increasing the supply while holding the price of use far below the market simply leads to chronic and aggravated congestion. It is like a dog chasing a mechanical rabbit. (Rothbard 1973, 213)

> [O]n the *market*, people are continually choosing between (usually) lower-priced but more crowded conditions, and more expensive, less congested alternatives. They do this in their daily choices to patronize, or not, a crowded fast food chain, a bargain sale at a local department store which they expect will attract large crowds, etc. The problem with our road network, in this regard, is that there is no functioning market in which the consumer can make his preferences known: there are no congested but cheaper highways, competing alongside more expensive but emptier ones. (Block 1980, 305)

[33] A similar statement appears in Peterson (1950, 200).

> Road pricing is a simple concept that extends the common practice that is virtually ubiquitous in every other sector of a market economy whereby prices are used to reflect scarcity, and to allocate resources to those that can best use them. In most places road space, even in such supposedly market orientated societies as the U.S. is in actuality allocated in a manner more akin to the general practices employed in pre-1989 communist Russia, namely by waiting in queues and lines. (Button 2004, 3)

> Are externality charges unfairly redistributive? They are not aimed at poor people but at voluntary activities: if you decide to stop causing trouble for others, you don't have to pay an externality charge. It is true that the rich can afford to drive more than the poor, but it is just as true that the rich can afford to eat more than the poor. This is unfair too, but if you accept the workings of the price system for typical goods like food, why not road space or clean air? We recognize that food, clothes, and houses cannot be free or we would quickly run out of them. It is because roads are free that we have run out of spare road space. (Harford 2006, 88-89 [writing for a general audience])

Marginal-cost pricing of roads

The tension between SRMC pricing to induce "efficient" usage of roads, and AC pricing to finance them has been evident since Dupuit and Pigou.[34] The choice matters to the case for road pricing inasmuch as marginal-cost pricing requires differentiation of charges with respect to space, time and vehicle characteristics, and thus calls for finer pricing instruments than does average-cost pricing. A uniform vehicle registration

[34] An early example is Clark (1923, 304-305) who discusses "the paradox of overhead cost" and the conflict between charging traffic enough to cover the overhead and the typically minimal charges (in his era) required to cover the marginal costs of their usage. This lesson is also central in Coase (1946) who advocated second-degree price discrimination. By contrast, Ramsey's (1927) solution for a set of products entailed third-degree price discrimination, and Ramsey pricing has been the more prominent as a policy candidate in both the peak-load pricing and road pricing literatures.

tax might recover costs at the aggregate level. The fuel tax could do so, too, and it would be more efficient and equitable than registration fees since payment of fuel taxes varies with amount traveled.[35]

The attitudes of some economists towards marginal-cost pricing of transportation have, no doubt, been influenced by the marginal-cost pricing controversy that was debated at length by Coase (1946) and others in the early and middle part of the twentieth century. Blaug (1985) reviews the "tortured history" of marginal-cost pricing (MCP)[36] and three passages are worth quoting:

> MCP is one of those orthodox doctrines that has been continually criticized and rejected by experts in the field of public utility pricing but nevertheless remains part and parcel of the corpus of received economic ideas. Even now, the precise status of the concept is a matter of frequent misunderstanding. (Blaug 1985, 16)

> The case for MCP, or, as we should now say, the case for making MCP a point of departure for a set of optimal prices, stems basically from the fundamental conditions for Pareto-optimal efficiency; and, of course, Pareto optimality is defined only with reference to a particular distribution of income or, rather, resource endowments. If we are unwilling to divorce efficiency from equity, at least for the sake of argument, neither the concept of MCP nor that of optimal deviations from MCP makes any sense . . . it is the willingness to analyse efficiency arguments apart from problems of income distribution that divides the advocates from the critics of MC(Blaug 1985, 25)

> MCP requires empirical judgements on a product-by-product basis about market structure, indivisibilities, externalities and elasticities of demand and supply; in short, it is a systematic check-list of what to look for in pricing a public service. It does not, therefore, furnish any

[35] Tolls can be used to internalize external costs besides congestion and road damages, but the case for doing so is not as clear. Greenhouse gas emissions, for example, can be effectively targeted with a carbon tax.

[36] See also Ekelund and Hébert (1999, Ch. 7) in the context of transportation.

> simple pronouncements about public pricing, except perhaps that public enterprises should not necessarily be expected to break even and that almost any pricing rule is better than average cost pricing. (Blaug 1985, 29)

The second passage from Blaug highlights the importance of attitudes towards equity, which will be covered later in this section. And the third passage points to difficulties in implementing MCP which will also be reviewed.

The competing demands of allocative efficiency, cost recovery, adherence to the user-pays principle, and equity were clearly identified by Beckmann, McGuire and Winsten (1956):

> From the point of view of the best overall expansion and utilization of a network, the constraint that each toll road should pay for itself must be dropped. Instead the problem becomes one of finding the most 'equitable' allocation of road costs to sources of tax revenue. To what extent should finance be sought in the form of taxes on vehicle ownership at particular locations on gasoline consumption (and hence total mileage, approximately) and on the use of particular roads (tolls)? (Beckmann et al. 1956, 5.12)

> While a case can be made for some use of general funds for this purpose, since an efficient road system contributes to the general welfare, the intensity of communication, and the speed of emergency help, etc., it would seem to be a point of justice that the bulk of the money should come from the road users in a form connected with road use. This would leave a considerable share to both general vehicle and gasoline taxes. The optimal apportionment—optimal, that is with respect to the combined standards of equity and efficiency—poses an interesting problem which will be the subject of discussion for a long time to come. (Beckmann et al. 1956, 5.14)

The last sentence is particularly prescient since the 'discussion' has gone on for half a century.

While most economists who wrote about road pricing in the mid-twentieth century were in favour of SRMC, there were notable exceptions.

Meyer, Peck, Stenason and Zwick (1959) focused on infrastructure costs rather than usage costs, and advocated an *incremental-cost pricing scheme* which effectively amounted to some hybrid of LRMC and AC pricing. However, their arguments (see 69-72) suggest some vacillation. Six years later, Meyer, Kain and Wohl (1965, Ch. 13) also appear undecided, but they too came down in favour of average-cost pricing:

> In essence, to adhere to marginal cost pricing in circumstances where marginal costs exceed average costs because of congestion amounts to imposing user charges on the basis of some elusive social cost concept rather than the cost of the physical resources consumed or used. Specifically, price discrimination based upon social or congestion cost concepts is quite different from price discrimination based upon differential resource costs created, say, by use of the facility during a peak period when extra and expensive capacity is needed to meet additional demands. Differential peak period charges to cover these additional capacity costs have a much sounder basis in economic theory and perhaps also in common equity as well. (Meyer, Kain and Wohl 1965, 339)

In her book on freight transportation, Friedlaender (1969, 130-135) favours basing rates on long-run costs, but is undecided whether the appropriate long-run costs are marginal or average. Smith (1975) reviews the arguments for and against SRMC pricing and AC pricing, and concludes:

> While there is a set of points upon which economists can agree, the weights given to the various points differ tremendously. The result is that reasonable men come to different conclusions when selecting the optimal pricing strategy for highways. (Smith 1975, 452)

> And Walters himself remarks: "The view that the user should pay for all the costs of the roads is strongly entrenched among many practicing transport economists" (Walters 1968, 4, paragraph 10).

These passages reflect the doubts that even leading transport economists have had about SRMC pricing. And differences of opinion still remain. This is evident from the exchange between Rothengatter (2003), who questions the practical value of SRMC pricing on various grounds, and Nash (2003) who defends SRMC pricing principles and describes progress in applying them in recent European research projects.

Importance of road pricing for investment decisions

Like Adam Smith, several economists have pointed out that implementation of road pricing would make the evaluation of highway investments both easier and more likely to be correct.

> There are . . . two aspects to the highway problem, which should be distinguished. There is, first, the short-run task of rationing the existing plant. When some semblance of efficiency has been attained in the utilization of the highways already built, the second task—that of building the optimum size plant—can be tackled more effectively and with less likelihood of over- or under-investment. (Meiburg 1963, 656)

> It will be said that it is costly to design and collect the correct charges. And because some grossness is necessarily encountered, we too lazily settle for clichés about the inefficiency involved in overpricing someone out of the market. Instead, we should be concerned about the inefficiency caused by not having the faintest idea about the true social value of the highway network in the first place and merely expanding it whenever serious damage or crowds are encountered, or the highway lobbies exert their pressure. . . . Certainly our highways are crowded. But until we attempt to price rationally, we have no way of knowing whether the market places a value on the additions to our road network that exceeds the cost of providing them. (Abouchar 1987, 53)

One longstanding question in the literature is whether the profitability of a road provides a reliable indicator whether or not it should be expanded. Roth (1967, Ch. 6) argued that roads on which revenues exceed costs should be expanded, and roads on which the revenues fall short should be contracted (or left to depreciate). In light of the cost recovery theorem, this rule is formally sound if there are constant returns to scale in investment and in usage, and first-best-optimality conditions apply elsewhere.[37] But absent these conditions, the profitability rule can be biased in either direction:

> Requiring each project to pay its own way may be the only way of making absolutely sure that the community does not persist in investing in uneconomical projects; but to adopt a policy that results in a substantial bias against undertaking increasing-return projects seems a rather costly method of insuring that errors in the other direction are avoided. (Vickrey 1948, 230)[38]

> [T]he fact that a proposed toll road may be self-supporting provides no automatic justification for its construction in boom periods. What is required in each case is an analysis of the costs and benefits vis-à-vis the costs and benefits from alternative highway policies. (Netzer 1952, 119)

Day (1998) makes the further point that environmental and other external costs should also be included in a project evaluation:

> [I]f road users are prepared to pay a price for the use of roads that is greater than the costs of providing additional road space (*including all the costs, externalities, land costs, a sensible measure of the costs of disturbing any areas with special wildlife and all the other genuine costs which can be identified*) then the additional road space should be built, and as in any other economic activity, the charge for the use of the new

[37] First-best conditions must hold since otherwise an investment will have spillover effects on traffic flows elsewhere on the network where usage is not efficiently priced. But if the whole network is efficiently priced, the envelope theorem applies, and evaluation of the costs and benefits of investment can be limited to the road link in question.

[38] Vickrey's stance contrasts with that of Coase (1946) espoused two years earlier.

> facility should be sufficient to finance its cost. (Day 1998, 7 [italics added])

Practical feasibility of road pricing

Some economists, such as Sharp (1966), have opposed widespread implementation of congestion pricing because of practical difficulties. There are several practical considerations.

Technology and administration costs

The costs of tolling roads were a major concern in the 1950s and 1960s; for example, the Smeed Report duly addressed such costs. Foster (1963, 245) expressed unease about the need for frequent and costly collection booths. Such concerns have faded, however, since the advent of affordable and reliable Electronic Toll Collection systems, which permit tolls to be varied by location, time of day and various vehicle characteristics. But for roads with low traffic volumes, the costs of tolling are inhibitive, and some economists still perceive them to be a problem generally.

> An important issue is how road users should be charged for road use. Many proposals have been made and some have recently been tried: tolls collected at toll booths, perhaps electronically; electronic metering of road use with bills sent to users at the end of the month; and fuel taxes. Each has its advantages and disadvantages; accuracy, capital costs and collection costs (both those of the collecting agency and those imposed on users) are the important ones. In the U.S., at least, an important issue would be illegal behavior to avoid fees. If users refused to pay their road use bills at the end of the month, the police would be forced to be collection agents; some people would demolish electronic gear in roads. In addition, users would become demoralized if there were many errors in charges in billing. Any sophisticated system would be several times more complex than the most complex activity of U.S. local governments. (Mills 1998, 78-79)

In addition to toll collection and billing, costs are incurred in assembling and implementing supplementary measures to make road pricing acceptable to the public. One such scheme, proposed by two engineers, Kockelman and Kalmanje (2005), and based on the idea of tradable emissions permits, is called "Credit-Based Congestion Pricing" (CBCP). CBCP entails congestion pricing on a network of urban highways. Residents of a prescribed area are each granted a monthly allowance of travel credits, and those who drive less than average can save the credit for future travel or exchange it for cash. In February 2004, Gulipalli, Kalmanje and Kockelman (2005) mailed questionnaires to transport economists and other professionals to assess their attitudes towards CBCP. The transport economists expressed general support for the scheme, but many were concerned about the costs of administering a credit system to a wide population.[39]

Pricing only selected roads

Except possibly for future satellite-based road pricing systems, it will never be practical to impose tolls on every road, and all schemes now in operation are limited either to downtown areas or selected highways. The area-based schemes either charge vehicles for moving within a charge area (London) or for crossing a toll ring (some Norwegian cities, Stockholm). These schemes catch only a fraction of trips, and the charges do not vary smoothly with distance. Some economists have objected to the crudeness of their design:

> The greatest growth in traffic in the UK over the last 10 years has been on motorways and the interurban road network. Should we be concentrating on toll road pricing on these routes rather than on city centres? Cities are themselves changing and the congestion problem may be migrating from the centre to the suburbs. The dynamic of the city is such that when road pricing becomes a reality, it is no longer a necessity and may accelerate the unsustainable city. (Banister 2002, 7)

[39] Attitudes of the economists towards the equity aspects of CBCP are noted later.

Shortly after London's congestion charge was launched, a similar cordon scheme was proposed for Sydney's central business district. Hensher (2003) expressed his misgivings:

> Although I support the notion of efficient pricing, this particular proposal is in danger of missing the mark …. The success of the London initiative as the first congestion charging program in a major European city is important for Australia in demonstrating the political feasibility of pricing …. What we must ensure however is that any congestion charging system is not selected for the convenience of an appealing cordon such as the CBD, but for broader systemwide efficiencies. (Hensher 2003)

Another drawback of both area-based charges and highway tolls is that motorists may divert onto toll-free routes, resulting in displacement rather than suppression of congestion and other external effects. However, Keeler and Small (1977, 23) downplay concern that tolling urban expressways will cause diversion to arterial streets because expressways (if well designed) will increase throughput and offer a much better quality of service.

Difficulties in computing optimal tolls

Beckmann, McGuire and Winsten (1956) showed that it is fairly straightforward to compute first-best tolls on a road network. But if tolls are restricted for any reason (e.g. only some links can be tolled, or tolls cannot vary over time) the problem becomes one of second-best. And the difficulties of computing second-best tolls have proved to be rather formidable.[40] The attitudes of economists vary. Most seem willing to support simple or crude schemes, particularly those that stand a reasonable chance of being implemented.

[40] This is true even for the prototypical Pigou-Knight two-parallel-roads network that has been analyzed extensively by economists since Lévy-Lambert (1968) and Marchand (1968). For more recent and general work see Verhoef (2000a and 2000b). Gómez-Ibáñez (1992, 354) is somewhat an outlier in considering the computational problem "manageable (and largely of esoteric interest to transport economists)."

> A comprehensive and highly differentiated system of tolls would be required to secure the 'ideal' pricing structure. Obviously, such a system would be completely unworkable from an administrative point of view, and would be uneconomic besides …. But with relatively little change or modification it appears that the system of highway user taxation now employed could be made to approach one which would achieve efficient operation of the existing highway structure. (Buchanan 1952, 102)

> As to over-simplification, I make no excuses for this. It seems to me that the complications that arise from the attempt to plan roads without a pricing mechanism are so great that over-simplification might even be welcome … If we want to obtain practical results in a real world we must use methods that can be readily understood by the civil servant and the local councillor. (Roth 1967, 13)

> [E]stimating the size of an optimal congestion tax is a considerably more complex matter than was recognised in some of the earlier literature, involving, as it does, a range of variables some of which can only be determined with a degree of uncertainty. This does not necessarily, however, weaken the case for using pricing methods to control congestion. The problem essentially stems from the difficulty of determining the optimal position, and that is a problem which would arise whatever regulatory instrument was envisaged as a means of achieving it. (Else 1986, 104)

> The best way to ration roads whose capacity cannot be increased is to charge a market price for the right to use them. Charges should be set just high enough to hold traffic down to levels that can move freely. Finding that optimum is less of a subjective exercise that it sounds: above a certain critical rate of flow, road traffic seizes up rapidly. The principle is equivalent to that of admission to a theater: it is self-defeating to let in so many people that everybody's view is spoilt, so sensible theater managers set prices that fill them to just below the point of discomfort. (The Economist 1989, 11)

> [I]t is not straightforward to calculate socially optimal congestion tolls by road sections and by time of day or day of week. However, since congestion consumes enormous amounts of a valuable resource (people's time), a reasonable approximation, which can be done quite easily, is likely to be better than completely dismissing this sound pricing method. (Gillen 1997, 211)

> Determined leadership using the available technology can deliver real benefits in the areas where they are most needed, before congestion becomes chronic with all the consequences for the quality of life and business efficiency, and thus international competitiveness. Doing something now, or very soon, is likely to serve the community better than deferring action until we can do it 'better'. . . Singapore, Austria and Switzerland reinforce the London lesson, demonstrating what can be achieved if sights are not set too high. 'Good enough' *can* be the basis of good policy. (Richards 2005, 287-8)

> Nobody knows the cheapest way of solving our traffic problems—yet. But externality pricing brings pollution, congestion, and the rest inside the world of truth, which markets create for us. As long as individuals have to face the truth, or at least our best estimate of the costs of their actions, they will find a way to reduce those costs. The longer they have to respond, the more surprising and innovative the responses can be. . . . The attractive thing about externality pricing is that it attacks the problem but makes no assumption about the solution. (Harford 2006, 95, 97)

Still, some economists have recently underscored the potential weaknesses of second-best schemes, or identified visceral resistance to them:

> [T]he amount of information required to apply a policy instrument to best advantage increases with the 'imperfectness' of this instrument. For the case considered

> here, this information includes the distribution of values of time and the demand elasticities of users having different values of time. Thus, second-best policies require considerable sophistication in order to achieve their theoretical benefits. (Verhoef and Small 2004, 154)

> [W]hen local politicians ask for low-price solutions, they are often abandoned by the scientific community. Engineers promote the most sophisticated technology available and show little enthusiasm for potential customers asking for low-cost technologies. Economists often do not dare to propose simple rule-of-thumb policies. Although such policies might reap huge welfare gains, they often fail to meet the professional standard of providing optimal or at least almost ideal solutions to actual economic problems. (Arnott, Rave and Schob 2005, 188)

As Verhoef and Small (2004) point out, the information needed to compute second-best tolls can be quite demanding. But, as noted in the previous section, even first-best tolling in the basic model requires information on demand and cost curves. To circumvent this problem many economists have suggested that tolls be set iteratively by a process of trial and error.[41] Setting tolls by trial and error was apparently anathema to the engineer St. Clair (1964) when he wrote "The user might experience some discomfort during this experimental period, but then, so does the guinea pig" (85). But periodic adjustment of tolls is now official policy for Singapore's electronic road-pricing system where tolls are adjusted every three months to maintain average vehicle speeds within a prescribed range (Singapore Land Transport Authority). Sandholm (2002, 2005, 2006) formalizes this idea by modeling an evolutionary game in which individuals adjust their travel decisions sluggishly in response to incentives, and the planner sets *variable* tolls iteratively with only limited knowledge about the individuals' preferences and behavior. As Sandholm describes it:

[41] See, for example, Hotelling (1938, 269), Friedman and Boorstin (1951, 238), Beckmann, McGuire and Winsten (1956, 19), Walters (1961, 697), Beesley and Roth (1962-63, 188,193), UK Ministry of Transport (1964, 35, 36), Roth (1967, 41), Vickrey (1967, 127; 1993, 4) and Beesley (1973, 229, 241-2).

> Variable price schemes can be viewed as generalizations of marginal-cost pricing. 'Marginal-cost pricing' usually refers to an equilibrium phenomenon: that by making agents pay for the externalities they create in equilibrium, one can guarantee the efficiency of equilibrium play. In contrast, variable price schemes set prices appropriately both in and out of equilibrium. (Sandholm 2002, 670)
>
> [V]ariable pricing enables a planner to ensure that efficient behaviour is an equilibrium without knowing what efficient behaviour will turn out to be. (Sandholm 2005, 887)

Complexity for users

As noted earlier a majority of economists support the principle of SRMC pricing which dictates that prices match the costs imposed by drivers as closely as practicable. Vickrey was a particularly strong proponent not only of varying prices with predictable variations in costs, but also responsively to unanticipated fluctuations. In defence of this policy he wrote:

> It might be thought that imposing charges determined after the user has very largely committed himself to the trip and the route would have little effect in improving efficiency of utilization. In this case, however, the larger part of traffic contributing to rush hour congestion is repeat traffic that tends to follow the same route at the same time day after day, and an additional portion of the traffic is a sufficiently frequent and regular user to become fairly familiar with the conditions to be expected at various points and times. Consequently, even though the imposition of a charge may come too late to affect the trip for which it is imposed, it will affect subsequent trips of a similar character, and thus is conducive to efficient utilization. (Vickrey 1971, 345)

Vickrey offers one explanation for traditional resistance to responsive pricing:

> [T]he main difficulty with responsive pricing is likely to be not mechanical or economic, but political. The medieval notion of the just price as an ethical norm, with its implication that the price of a commodity or service that is nominally in some sense the same should not vary according to the circumstances of the moment, has a strong appeal even today. (Vickrey 1971, 346)

Despite his enthusiasm for responsive pricing, Vickrey acknowledged even in his early writings the limitations of SRMC pricing in its pure form.

> [I]t is at least doubtful whether any advocate of marginal-cost pricing has ever seriously proposed that prices should slavishly follow marginal cost in every detail, without making some allowance for administrative costs involved in such detailed rate structures and for the fact that, beyond a certain point, the consumer may become so confused that the more intricate rate schedule would cease to function effectively as a guide to consumer choice, thus losing its *raison d'être*. (Vickrey 1948, 233)

Many other economists have advised against varying tolls responsively[42], or even in short time increments according to a predictable schedule.

> If the price system is complicated road users will probably find simple 'rule of thumb' methods to tell them approximately what the average prices are and roughly what the prices of particular journeys are likely to be, and they will act accordingly. If this is so the complicated system may be no more efficient than a simpler system.[43] (UK Ministry of Transport 1964, 48)

[42] To date, responsive pricing has been applied only on two High Occupancy Toll facilities in the US: Interstate 15 in San Diego, and Interstate 394 in Minneapolis.

[43] Intriguingly, Michael Beesley, a member of the Smeed Committee, appeared to back away from this position five years later: "Two other desiderata, which the Committee thought were important now seem less so. These were that 'prices should be stable and readily ascertainable by road users before they embark on a journey'; and that 'the incidence of the system on individual road users should be accepted as fair'. The first merely turns on experience and information; even now, no one can predict with certainty, when proposing to

> The toll itself must also be comprehensible. Thus a simple peak/off-peak pricing differential (marketed as an off-peak discount) is much preferred to a congestion-based variable toll. Variable tolls are conceptually elegant and technically feasible, but potentially confusing to the consumer. (Giuliano 1992, 354-355) [44]

> [S]imultaneous pricing is subject to the same sort of objection as retroactive pricing[45], because motorists would not know the price until they were committed to their trip. From an economic point of view, simultaneous pricing is inefficient because it forces consumers to make decisions on the strength of uncertain knowledge of the prices involved. Such uncertainty would probably make it unacceptable to the public. (Thomson 1998, 99)

> Without questioning the legitimacy of marginal cost pricing, the problem is that is has proved difficult to implement. In today's technologically advanced world the calculation of instant marginal cost pricing may not be very difficult to envisage. Its cost effectiveness, however, would be dubious and, most importantly, the transparency of such a system would be at least arguable, as drivers would not know the congestion charge they would be required to pay before starting their journey. Marginal cost pricing would require highly differentiated pricing systems in time

make a trip what congestion will be: certainly under road pricing expectations about congestion, and therefore prices, will be no less certain." (Beesley 1973, 234)

[44] Genevieve Giuliano does not meet the definition of an economist adopted for this review as "someone with a postgraduate degree in economics or a job with a title of economist such as a teaching or research position at a university economics department." She has a Ph.D. in Social Sciences, U. of California, Irvine School of Policy, Planning & Development, and is currently a Professor in the School of Policy, Planning, and Development, University of Southern California (http://www.usc.edu/schools/sppd/faculty/detail.php?id=11, accessed May 3, 2006). Her background is "in geography, economics and political science" (http://152.122.41.184/NYSMPOs/colloquy_bios.asp, accessed May 3, 2006). Given the overlap of her education and background with economics and her extensive publication record, I considered it fitting to include her in the survey.

[45] As Thomson defines the term "simultaneous pricing" is synonymous with responsive pricing. "Retroactive pricing" entails setting prices after a trip is completed.

> and space, which would be expensive to provide and confusing to users. (Santos 2004, 346)

Equity

Welfare-distributional considerations were addressed early on in the public utility pricing literature (Feldstein 1972). But equity was largely ignored, or dismissed, by both the pioneers and the mid-century writers on road pricing. As Beesley remarked:

> Undoubtedly the main opposition to road pricing arises from doubts as to its effects on individual's welfare. To some extent, exponents have themselves to blame for what has often been a hostile reception. Partly this is because in early expositions the case was drawn up in such a way as to eliminate the problem of inter-personal comparisons of gains and losses. The Knight-Pigon [sic] example postulated a 'narrow but good' road competing with an inferior one of infinite capacity. No one could be made worse off by tolls on the narrow but good road when congestion arose on it, because total costs rose to the level of those on the 'bad road', which acted as a perfect substitute. . . . The Smeed Committee's own calculations of 'benefits' from road pricing treated road pricing as a 'transfer payment' which represents no real cost to the community. (Beesley 1973, 279-280)

Some of the first economists to consider the welfare distributional effects of road pricing, such as Layard (1977) and Glazer (1981), concluded that tolls are regressive because those who benefit the most have the highest values of time (VOT) and VOT is positively correlated with income. Foster (1974), however, pointed out that the welfare impacts of road pricing depend not only on changes in travel time and out-of-pocket costs, but also on any improvements in public transport service, changes in the geographical incidence of emissions, accidents and other external transport costs, and so on. Economists came to appreciate the difficulties in assessing the welfare impacts as well as the practical impossibility of designing a tolling scheme that leaves everyone better off. Partly for this

reason, many economists have downgraded equity as a primary consideration for road pricing.

> Disputes over the distribution of the national income can be handled much more reasonably if they are brought into the open in discussions of rates of income tax and other deliberately income-redistributing measures. Such considerations can be excluded from rate-fixing problems only by setting rates at marginal cost. (Vickrey 1948, 236)

> Section 210 of the Highway Revenue Act of 1956 suggests 'equity' to be the basic desideratum of Congress in allocating the burden of highway finance. Unfortunately, 'equity' is not an operational concept. The word means 'fairness' and 'justice,' which are terms that do not have universally accepted operational definitions. (Mohring and Harwitz 1962, 87)

> That they are ill-fed, ill-clothed, ill-housed, and, perhaps, ill-transported is really not the basic problem that poor people have. Their fundamental problem is, rather, that they are poor. If we are genuinely concerned with how road pricing would affect their welfare, we should give them cash or marketable road scholarships. It would be the height of folly, however, to subsidize all of our private-passenger-vehicle road use in the supposed interests of helping them. (Mohring and Anderson 1994, 34)

Foster (1974, 1975) and Richardson (1974, 1975) exchanged views on the importance of equity:

> [S]ince there is a strong case for road pricing on efficiency grounds, whereas the equity arguments are murky, the issue should be decided in the light of efficiency, goal achievement and political feasibility. We should not try to make a probably unpopular but effective measure more palatable by resorting to specious social arguments. The question of the equity of the tax among road users as a whole cannot be resolved, and hence is not critical to decisions on road pricing. (Richardson 1974, 84)

> (Of course this is not to say that the state could so determine net distributional impact for each and every person as to turn the change into a Pareto improvement or any other exact distribution of the net gains as between *individuals*.) That introduction of road pricing is a potential Pareto improvement is the single most important proposition relevant to its net distributional impact.[46] (Foster 1975, 186 [after noting that the state influences the welfare effects of road pricing by choosing how to allocate revenues])

> I also support direct road pricing, though on grounds that do not require buttressing with 'murky' equity arguments, and I agree that the revenue can be used as an instrument for compensation. . . . I would like to endorse Foster's statement that aggregate measures of income distribution are of little value and his suggestion for a 'net distributional impact' approach to the evaluation of transport policy decisions. If equity objectives are relevant to transport policy, and I think they are, measures that minimize unfavourable distributional impacts on the poor merit serious attention. (Richardson 1975, 188 [now showing rather more concern for equity])

After accepting the potential Pareto improvement as a criterion for project evaluation, Glaister (1981) addresses the equity of road pricing:

> The 'unfairness' of peak pricing rests on the premise that constant prices are 'fair'. But it is neither fair nor sensible to encourage the poor or anybody else to use a facility which costs society (or travellers at other times) more to provide than the benefits they derive from it. On the other hand, it must be admitted that the imperfect nature of taxation and compensation systems in practice means that inevitably some individuals will in fact be made worse off.

[46] This echoes the passage from Blaug (1985, 25) quoted above in the section "Marginal-cost pricing of roads".

> This disadvantage has to be set against the welfare losses due to economic inefficiency. (Glaister 1981, 69)

Not all economists, to be sure, disregard equity as a desideratum for road pricing. For example, the Gulipalli et al. (2005) survey of attitudes towards Credit Based Congestion Pricing found that some transport economists considered inequitable a distribution of credits to everyone with a driver's license, regardless of how much they would use toll roads.

A number of economists have examined the relationship between efficiency and equity inherent in the design of road pricing schemes. Most have concluded that there is a tradeoff between them. Some have examined the efficiency-equity tradeoff that has been implicitly accepted in the decision to toll only some lanes of the High Occupancy Toll (HOT) lane facilities in the US. Some consider the tradeoff to be reasonable (e.g. Giuliano 1992, 354; Richardson and Bae 1998, 259). Verhoef and Small (2004) show using a simulation model that the tradeoff depends critically on how total road capacity is allocated between the toll lanes and the toll-free lanes.

Some writers claim that road pricing is no more regressive than other ways of paying for roads.

> The existing system of road financing is triply regressive. Taxpayers generally, a poorer class, support road users generally, a richer class, with municipal services like fire, police, courts, ambulances, and emergency rooms benefiting road users. Road users generally, a poorer class, support peak-hour road users by paying with their fuel taxes for roads generally sized to accommodate peak-hour users. And five out of the six taxes supporting the existing highway system are themselves regressive. (Giuliano 1994, 260)[47]

To the extent that tolls are no more regressive than other taxes, using toll revenues to reduce other taxes in a revenue-neutral manner would not harm equity.[48] This prospect is especially relevant for fuel taxes, as

[47] As quoted by Elliott (2000, 10).
[48] Cameron (1994) makes this argument forcefully for the case of gasoline taxes.

proposals for area-wide tolling in Britain and Oregon would entail reductions—or even elimination—of taxes on fuel.[49]

Allocation of road-usage revenues and earmarking

The allocation of revenues from road-user charges affects the welfare-distributional effects of road pricing as well as its public acceptability. Indeed, since charge revenues may be a large multiple of the efficiency gains from road pricing (cf. Figure 2) the allotment of revenues may be the dominant factor. According to some public finance textbooks, tax revenues should not be locked into any particular expenditure pattern because priorities can change in unforeseen ways. Nevertheless, designating or earmarking [50] revenues for particular purposes is common practice in transportation as well as other sectors and it has been justified on various second-best grounds. Economists have taken different stances towards earmarking of road-usage charge revenues. At least four camps can be identified: (a) those against earmarking, (b) those in favour of revenue neutrality (i.e. offsetting new road-user charge revenues with reductions in other charges), (c) those in favour of earmarking generally, and (d) those in favour of particular earmarked allocations.

(a) Against earmarking

> Plowing congestion tolls back into road improvements is not necessarily efficient. Presume zero population and travel growth, infinitely durable roads, and an optimally

[49] The gasoline tax is widely considered a regressive tax. Most empirical studies that draw this conclusion use data on household fuel tax outlays as a proportion of annual income. Poterba (1991) argues that because of transitory shocks and life-cycle variations in income, annual income is a poor measure of economic well-being. Using US data on annual household expenditure instead of annual income, he finds that the gasoline tax is actually slightly progressive over the bottom half of the income distribution. However, Chernick and Reschovsky (1997) criticize this approach on the grounds that annual expenditure is a poor proxy for lifetime income. To construct an intermediate-run measure of the gasoline tax burden they use eleven years of panel data on US household income and gasoline consumption. The tax burden computed this way is only slightly less regressive than the burden based on annual incomes. These disparate results illustrate the difficulties in using household incomes to assess the equity effects of gasoline taxes, tolls or other road user charges.

[50] Commonly referred to in the UK as hypothecation or ring-fencing.

> designed road network. Marginal-cost tolls would then function as a normal return on the resources that society has invested in its road network. Efficiency would dictate using road-user tolls just as any other source of government revenues. Efficiency would not dictate spending these revenues on road improvements. (Mohring 1991)
>
> It would be undesirable to return most of the proceeds to motorists by investing them in highway or transit improvement since the surpluses often may be more of an accounting artefact than a sensible signal that added capacity is needed. (Gómez-Ibáñez 1992, 358)
>
> "[I]t makes no more sense than any other scheme to hypothecate tax revenues for specific sorts of spending: if a programme is sufficiently worthy, it should be financed regardless of where the tax-money comes from; if not, the money should be spent elsewhere. A clever transport minister, in his speech proposing road pricing, will nevertheless want to make voters feel that as travelers they would benefit:
>
> 'A good transport system is the handmaiden of a healthy economy. It also allows people an essential personal freedom—to travel as they wish. The government wants to meet the demand for more travel, including car travel, not to suppress it. The money raised from city drivers will not be earmarked for particular transport projects, but you can be sure that it will make it easier for the nation to afford new road and rail links generally.'
>
> Expect such words to be intoned in many great cities before this century is out." (*The Economist* 1989, 12)[51]

[51] *The Economist* seems to take a nuanced stance by allowing that at least some revenues should be devoted to transportation as a whole.

(b) In favor of revenue neutrality

> [T]he adoption of a pricing scheme to allocate road space need not necessarily raise the total tax burden. Our argument is concerned with the *distribution* of the tax burden, not with its absolute level, and we advocate congestion taxes as a substitute for—not as an addition to—license, purchase and fuel taxes. The adoption of such a policy does require judgments to be made of the gains and losses to different groups of road users, but if the total of tax were to remain the same, there is no doubt that a more sensible allocation would benefit road users as a whole. (Beesley and Roth 1962-63, 196)

> The proposition that needs to be put to the public is that in exchange for the entire system of current road taxes (fuel taxes in excess of the rate of VAT, the special car purchase tax, and the licence fee), road-users will be charged according to their use of congested road space, at a rate which for the average road-user will be roughly the same. . . . As more than half the road-using population drives less than the average number of miles in congested areas, this should command majority support. (Newbery 1990, 31)

> I would propose . . . a much more drastic reduction in fuel tax than anyone at present envisages . . . in part counterbalanced by congestion taxes through the main congested conurbations. . . . Now is the time to reduce fuel taxes by more than the amount raised in congestion charges. (Walters 2002)

(c) In favor of earmarking generally

> The re-allocation of revenues outside the transport sector does not seem a good idea, and is likely to stimulate existing prejudices of road users as being one of the government's favourite 'cash cows'. (Verhoef, Nijkamp and Rietveld 1997, 272)

> Subject to further research, the idea of setting up a transportation (or road) fund to pursue marginal cost pricing in all its dimensions would enable us to satisfy the quintipartite principles of the World Bank's general guidelines for improving transport efficiency . . . to: (1) implement efficient pricing; (2) meet economic viability; (3) meet (to a considerable extent) financial viability; (4) achieve (some degree of) 'fairness' among beneficiaries; and (5) attain (somewhat) managerial efficiency of the public authority. (Hau 1998, 69)

(d) In favor of particular earmarking schemes

Earmarking has been gaining favour amongst advocates of road pricing, and a variety of schemes have been suggested. Goodwin (1989) proposed a "Rule of Three" whereby revenues would be allocated in equal parts to: (1) development and maintenance of new road infrastructure; (2) public transport; and (3) either to reducing the general tax burden or to increased spending. He acknowledges (p.496) that these proportions are arbitrary. But in Goodwin (1997) he argues that a scheme of this sort is essential to overcoming public opposition to road pricing:

> [R]oad pricing without explicit attention to the use of revenue streams is *inherently* unlikely to be able to command a consensus in its support. I treat this as an axiom of contemporary transport policy. (Goodwin 1997, 2-3)

Like Goodwin (1989), Small (1992) advocated a tripartite division of revenues although the goal of his scheme was less to spread benefits widely than to prevent opposition.[52] After an extensive literature review on acceptability of road pricing, Ison (2004, 74) concludes that hypothecation is "all-important". He further notes

> The research suggests that the largest proportion of the revenue generated from road user charging should be

[52] This assessment is drawn from King et al. (2006).

> utilised to improve public transport, particularly in the area in which the charge is introduced. (Ison 2004, 175)

From an efficiency standpoint, however, Button (2006) advises against earmarking to public transport:

> There would seem from the experiences of Norway and from some of the admittedly limited survey analysis conducted of road users in the UK, that if ring-fencing of revenues is done then there is at least a better case for ensuring that monies are devoted to road improvements where there is demonstrable demand, than to public transportation. Indeed, since public transportation often uses a common track there are potential synergies. (Button 2006, 239)

Most proposals for earmarking are based on the idea that road pricing schemes must create more winners than losers. King et al. (2006) argue to the contrary that gaining acceptability requires that the benefits from road pricing be concentrated, and the losses dispersed, so that the winners have enough at stake to overcome any opposition. To accomplish this they recommend that toll revenues be given to city governments where highways pass through.

> Congestion pricing cannot be sold as a policy that harms no one, nor even as a policy that helps everyone. It needs to be positioned as a policy that will help some particular group a lot. We believe that constituency can and should be the cities that host freeways. Congestion pricing in this scenario can benefit from the established strength of intergovernmental lobbyists, and can at the same time be a progressive force that compensates areas near freeways for the negative externalities they suffer. (King et al. 2006, concluding paragraph)

King et al. (2006) do not advocate that city governments spend toll revenues on any particular set of goods or services. Rather, their proposal is for earmarking revenues to particular *institutions*.

Finally, in their survey of the attitudes of transport economists and other professionals towards Credit-Based Congestion Pricing, Gulipalli et al. (2005) found some diversity of preferences among economists:

> [T]ransport economists were asked to rank a set of alternatives for uses of 'excess revenues'. Most wanted such revenues to go toward maintaining existing infrastructure and/or adding capacity. Next was development of alternative modes, such as transit. Those who strongly favored transit were not interested in reducing gas taxes—and vice versa. Some respondents suggested reducing general taxes via CBCP revenues. There was not much interest in using such revenues to improve air quality. (Gulipalli et al. 2005, 6)

In summary, economists have a range of opinions on how the revenues from road pricing should be used. A growing proportion (and perhaps now a majority) seems to favour earmarking revenues in some way. However, it may be that they support earmarking only reluctantly as a necessary concession for road pricing to move forward.[53]

In favor of road pricing, but not too hopeful

Again, concerns about political acceptability intermingle judgments of desirability. Some economists support road pricing but are less than sanguine that it will ever be implemented on a large scale. The following statements are listed in chronological order.

> The Hong Kong experiment has shown the rest of the world that the hardware can readily be designed for electronic road pricing. However, overcoming the political implementation problems is much more difficult. If they cannot be overcome, then electronic road pricing may forever sit unused on the economist's shelf. (Borins 1988, 44)

[53] If so this would be an example of the classic tension between the desirable and the politically acceptable noted in the introduction.

> After more than 35 years I am still involved in these arguments [about the virtues of SRMC pricing]. Although… the *principle* has been largely conceded, the applications are much more timorous and messy. . . . I do not retain a faith that, because it is so rational and liberal a solution, ultimately it will be widely adopted. (Walters 1988, 20)
>
> [I]t is unlikely that congestion pricing will be implemented to any significant extent in the U.S. Public scepticism regarding the effects of congestion pricing, resistance to high tolls, and pressures to divert toll revenues to new transportation facilities are barriers to effective congestion pricing programs. More likely are tolls on new capacity, tolls for specific classes of users, and other less direct and less complex auto pricing strategies. (Giuliano 1992, 335-6)[54]
>
> [S]ome transport economists and environmental planners are encouraged to believe that motorists soon may pay, in some places at least, the marginal costs of road use. The general public, however, is not attracted to tolls and congestion pricing for the same reasons that interest economists and environmentalists. The public and its elected representatives are primarily interested in tolling as a means to finance the expansion of facilities rather than as a means to manage existing facilities more wisely. As a result, tolling is not likely to be implemented in forms that economists or environmentalists would recommend. Tolling is more likely to be implemented in a piecemeal rather than a comprehensive fashion, primarily on new or expanded facilities. Tolls are also likely to be set at significantly lower levels than either economists or environmentalists would suggest. This combination of piecemeal tolling and low prices may seriously limit gains from tolling in many situations. (Gómez-Ibáñez 1992, 343)

[54] As quoted by Richardson and Bae (1998, 250).

It would be over-optimistic to expect the sophisticated charging schemes tested in the last decade, on a small scale, to be implemented soon on a nationwide scale. This is particularly true for the less developed countries (LDCs), where even standard charges are plagued by problems of tax administration and gross evasion. In the foreseeable future it is realistic to expect that policy makers will have to rely on the standard battery of taxation tools, foremost among which is the fuel tax. (Gronau 1994, 255)

We shall probably have to wait for particularly high levels of congestion to make tolls acceptable to automobile lobbies. The relevance of the theoretical argument is therefore insufficient: a whole array of socio-political parameters involving national temperaments interfere in the question of the acceptability of urban tolls for society. (Derycke 1998, 71)

It is probable that significant portions of an urban highway and road system cannot be subjected to congestion tolls. For example, many people reside on arterial streets and highways that carry a good deal of rush-hour traffic. It is not likely that any politician will ever suggest that people should pay a toll to drive on their own streets. In our view congestion tolls are not likely to be imposed on more than the limited access portion of the urban highway and road system. (McDonald, d'Ouville and Liu 1999, 234)

Contemporary citizens dislike taxes and distrust governments. Unless (a) urban travelers can be persuaded that Bangkok-type equilibrium prevails during peak periods or (b) losers can somehow be reimbursed from toll revenues in a fashion that does not distort their travel behavior appreciably away from the with-toll optimum or (c) detailed toll-revenue-expenditure program can be found that a substantial majority agrees justifies paying tolls, congestion-pricing packages will continue to be a very hard sell. (Mohring 1999, 198)

> The more sophisticated the scheme . . . the closer it is likely to get to the Pareto optimum solution. . . . However, what is the most efficient and 'economically pure' as a road user charging scheme is also likely to be less popular politically. There is thus a dilemma in that the more sophisticated scheme is more likely to change driver behaviour, whereas the less sophisticated is likely to be more acceptable. (Ison 2004, 17, fn. 7)

Downs (2004) runs hot and cold on road pricing during the course of his recent 455-page book. He considers demand-side measures to be generally more effective at reducing congestion than supply-side tactics, and also less costly (p334-5). But he believes that road pricing will never be widespread in the United States:

> As an economist, I favor market-based approaches whenever possible. However, their political feasibility has been restricted by the egalitarian American desire not to provide any relative advantage to high-income travelers versus low-or moderate-income ones. This desire is politically potent because of the high rate of automotive vehicle ownership in America. Moreover, households considering themselves in the low- and moderate-income category vastly outnumber those considering themselves to have high incomes. (Downs 2004, 327)[55]

In one particularly negative passage Downs writes:

> If most Americans clearly understood the alternatives, they would undoubtedly regard congestion as much better than rationing space during peak hours by using tolls on all major roadway lanes, or building vastly more road space to avoid such rationing altogether, or trying to expand public transit systems enough to absorb all those 'excess drivers' seeking to use the roads in private vehicles during peak hours. (Downs 2004, 11)[56]

[55] Similar statements are found on pp. 79, 161.

[56] An interesting twist on this is the view expressed by Calfee and Winston (1998, 96-97) that road pricing should be adopted in order to preclude the other policies: "Although our

Downs's overall conclusions appear to be: (a) congestion is here to stay (Ch. 1), (b) people should learn to adapt to it (13), and (c) people should find ways to make congested time more enjoyable (354).

ATTITUDES TOWARDS PRIVATE ROADS

For three reasons the attitudes of economists towards private roads are reviewed separately from their stance towards road pricing on public roads. First, a large majority of the literature on road pricing deals only with public roads. Second, many economists support road pricing on public roads but are against private roads, whereas for other economists private roads are the preferred alternative. Third, private roads raise a number of new policy issues such as the form of ownership or contract (design, finance, build, operate, transfer), the type and stringency of toll regulation, and the exercise of powers of eminent domain. Economists may oppose private roads not on principle, but rather because they believe that all these issues cannot be satisfactorily addressed in practice.

Background

Probably the most common reason for economists to oppose private roads is that they see roads as natural monopolies that confer private operators with market power to raise tolls above efficient levels. Several economists have shown that under conditions of perfect competition the market equilibrium will coincide with the social optimum (Buchanan 1956, Vickrey 1968, DeVany and Saving 1980). But it is typically argued that the scope for competition is in practice rather limited (e.g. Gómez-Ibáñez and Meyer 1993).

Two arguments in favor of private roads are often made. One is that private owners have a greater incentive than do public institutions to achieve productive efficiency and to seek innovative ways to cut costs

findings cast doubt that the (direct) net benefits from congestion tolls are large, policymakers should still pursue the policy, if only to head off other—very costly and far worse—approaches to reducing congestion."

and/or improve service quality. The other argument is that government agencies themselves may succumb to the temptation to boost revenues; by raising tolls above first-best (or second-best) levels, by restricting capacity, or (if compensated via fuel-tax receipts) by making roads rough in order to increase fuel consumption (Smith 1937, 686; Foster 1963, 252-3; Evans 1992, 234; Roth 1998, 13).

The economists surveyed here are organized into three groups: those that support private roads with little or no reservations, those who are ambivalent or cautious, and those who are opposed. Each group is sampled in roughly chronological order.

A. Supportive of private roads

As has been discussed, Dupuit and Knight appear to have been favorably disposed towards private roads, but their views are open to interpretation. The first clear and unequivocal support for private roads appears to be an essay by Milton Friedman and Daniel Boorstin (1951). Because of the article's perspicacity and scope it is worth summarizing. Friedman and Boorstin begin with the statement

> The building and maintenance of our highway is today almost exclusively a governmental operation. We have become so used to this that whenever the question arises how to solve our highway problems, we take it for granted that we are simply asking how we can improve government planning of roads and government financing of them. (Friedman and Boorstin 1951, 223)

Friedman and Boorstin identify three main obstacles to efficient operation of private roads: (1) the technical difficulty of charging for the use of roads; (2) monopoly power; and (3) the 'neighborhood effect' of roads; i.e. since residents benefit from the access provided by roads it is unfair for through-traffic to bear the full cost of maintenance. Friedman and Boorstin consider these obstacles least important for turnpikes. They recommend (232-233) turning turnpikes over to private enterprise and giving it free reign in setting tolls. To reduce unfair competition from toll-free public roads, the state would rebate to the firms the approximate fuel tax revenues paid by motorists while driving on their turnpikes.

Friedman and Boorstin see greater difficulties for ordinary inter-city roads in charging for use and containing monopoly power. As a means of charging they suggest (238), rather fancifully, painting radioactive material on the centre lines of roads, equipping vehicles with Geiger counters, and charging according to the radioactive intensity recorded at a rate dependent on the type of road. Should no practical means of charging be devised, they recommend that private enterprise be allowed to compete for traffic on the quality of maintenance and other services provided, with compensation paid by the state equal to the taxes collected on gasoline consumption (239). Finally, Friedman and Boorstin admit to being unable to devise a workable plan for private operation of urban roads.

Amongst the other supporters of private roads Rothbard (1973, 203-204, 214-215) describes how private road owners would impose congestion tolls that encourage travelers to respond in beneficial ways including carpooling, shifting to public transport, changing work hours or place of residence, etc. Profits would encourage firms to expand capacity, maintain safe operations and efficient police protection. Similar arguments are presented in Block (1979, 1980, 1996) and Cadin and Block (1997).

Fielding and Klein (1993) assume that highways will be franchised, rather than outright privatized: "We are optimistic about competitive bidding for highway franchises, but viability will depend on *how* the bidding is organised" (Fielding and Klein 1993, 114). They offer a mixed assessment of the need for toll regulation, voicing more concern about excessively high tolls during off-peak periods than during peak periods when reductions in congestion are desirable.

Several chapters by economists in the recent volume on private roads (Roth 2006) assess the advantages of the private sector in building, maintaining and operating roads. Three chapters will be mentioned here. Klein and Majewski (2006) review the history of turnpikes and other forms of private toll roads in the US. They document how private companies provided a higher quality service at a lower cost per mile than did comparable public roads, and succeeded in developing extensive networks of highways.

> Looking back, one might say that the American people ran an experiment: 100 years with extensive privately managed toll roads, and then another 100 years primarily of government managed 'freeways'. The historical record suggests that road provision is another case where the advantages of private ownership relative to government

> ownership, and of user-fees, relative to tax financing, apply. Learning from the mistakes of both epochs, Americans and people in other countries should embark on a new century of road provision. (Klein and Majewski 2006, 300)

Button (2006) and Foldvary (2006) provide more contemporary views:

> If roads were provided privately and competitively then the inefficient congestion problem would not exist. Road owners would price to maximize profits and in so doing would take into account the costs of congestion each driver imposes on other drivers. In addition the issue of net investment would also be taken care of as prices would accurately convey information about where and when to build new roads. (Button 2006, 226)

> [P]rivate streets are not merely economically feasible, but superior in efficiency and service in the financial and organizational context of the decentralized, competitive, and responsive private communities in which they would be provided. From a purely economic and ethical perspective, it is not private streets but governmental streets financed by forceful means that require justification and explanation. (Foldvary 2006, 323)

Klein (1998, 14) argues that the organizational advantages of the private sector for tolling roads also apply to the provision of urban transit services:

> Imagine the city streets and roads divided up into segments or small districts. Each separate unit would be under the control and management of a private entity. . . . Just as shopping malls allow free parking, street owners might make road access one of the gratis attractions to visitors, residents and businesses. Just as proprietary communities often provide minibus service gratis, the road-owner might provide free bus service. Alternatively, the road-owners might implement electronic road pricing. . . . The

> natural incentive is for the road owner to work with associations and agents that coordinate the interdependent parts of the road and transit system. In private industry, such standards for matters of technology, product design, product safety, and insurance emerge from voluntary machinations—both competitive and cooperative. We could expect the same for transit coordination. The natural incentive is for the road-owner to form contracts that will enhance his road as a place to shop, work, and reside.

An important point regarding cost recovery that comes out in the excerpt from Klein (2006) is that private owners can decide whether and how to charge explicitly for the use of their roads. In some cases it may be unnecessary to recover all costs through tolls. For example, in the case of a privately developed gated community the costs of local roads can be built into housing prices and homeowner association dues. Likewise, merchants can recover the cost of access roads to their establishments in the prices of the goods they sell.

B. Agnostic, ambivalent or cautious about private roads

Roth is included in this category because his early writings leaned against private roads. Over time, however, his attitude became more positive.

> [W]hile it is possible to envisage competition in the provision of roads connecting points at great distances apart—as occurred on the railways in the early days—it is not possible to envisage competition in the provision of access roads in towns and villages, for most places are served by one road only. A highway authority is in practice in a monopoly position. If any of its roads were to make large profits, we could not expect other road suppliers to rush in to fill the gap. If losses are made on some roads, there are no road suppliers to close them down and transfer their resources to other sector of the economy. (Roth 1967, 63)

> After the introduction of road pricing in Singapore, Walter Block was writing revolutionary articles suggesting that

> roads should be privatized, and warning that the imposition of congestion pricing on roads without including them in the market sector was unlikely to result in optimal solutions to transport problems (Block, 1979). It was this seed, planted by Walter, that eventually brought me to write this book. The question it asks is to what extent the concepts of ownership, free prices and voluntary exchange—concepts that govern the provision and allocating of scarce resources in free societies—can usefully be applied to roads. The book discusses the possibilities of public roads being privately provided, but its thrust is directed more at the *commercialization* than the *privatization* of roads. This is because I see the commercialization of roads . . . as a major objective in its own right, as well as a necessary step on the road to privatization. (Roth 1996, xix)

> On the one hand, where a public road is privately provided as a result of voluntary transactions — where, for example, no governmental powers are used to purchase land, and where providers of new roads are given no protection from competitors. In those circumstances, it is difficult to justify government interference in the rights of the owners to set any fees they please. On the other hand, where government powers are used to obtain land, or where a private supplier is given protection from competition, an arrangement to limit the profits of the enterprise would seem to be reasonable. (Roth 1996, 102)

In the introductory chapter to Roth (2006), Roth comes even closer to an endorsement of private roads by listing the advantages of privatization (11-13) and opening his chapter with the statement:

> Will Rogers is reported to have said that the way to end traffic congestion is to have the government build the cars and private industry the roads. The purpose of this book is to demonstrate that only the latter recommendation is necessary. (Roth 2006, 3)

In discussing private roads, David Friedman (1989 [1969], 72-74) describes the advantages of peak-period tolls and the feasibility of electronic tolling, but explains the challenges in guaranteeing road access for homeowners and defining legal rights and responsibilities with regard to access roads.

Beesley (1973) addresses the potential role for the private sector in financing new investment. In favour he remarks

> Its real merits, one suspects, may well be simply that of opening the road system to innovatory enterprise—new standards of safety, attractive driving conditions, more competitive tendering for contracts, the spread of new standards for construction and maintenance, etc. (Beesley 1973, 251)

But Beesley (1973) also identifies capacity indivisibilities as a constraint on competition, and expresses doubts that investments will be made at the best locations on road networks (251).

Beesley and Hensher (1990) adopt a stance similar to Beesley (1973). They offer a balanced and "speculative" (331) assessment of private sector involvement in ownership and operation of tolled facilities. In their concluding remarks they write

> The most important issues include the extent to which competition or price control provide the best mechanism for protecting the consumer, and aiding efficient supply of road space; the desirability of establishing independent regulation and rivalry amongst operating authorities, the need to have a clear understanding of the cost of capital and the sources or risk; and the extent to which turnkey or termination deals are really a desirable strategy. (Beesley and Hensher 1990, 340)[57]

The following passage by James Buchanan expresses doubts about the political acceptability of private roads:

[57] A similar attitude is expressed by Hensher and Puckett (2005, 382).

> Are the arguments of the economists more likely to find receptive audiences in the post-socialist politics of the 1990s and beyond? My skeptical public-choice instincts suggests a negative answer… Separated private ownership of most components of the road network remains a dream only for the most utopian libertarians. . . . But depoliticization via commercialization is both economically and politically realizable. (Buchanan 1996, xv)

Winston and Shirley (1997) consider privatization a serious possibility, but only in the longer run: "In the absence of accumulated empirical evidence, we believe it is premature to recommend privatization of U.S. highways" (Winston and Shirley 1997, 103). Their main fears are that government agencies will regulate tolls, "discourage efficient pricing schemes" and mismanage advanced navigation and road traffic control technologies. It may be that these modern concerns contribute to their less sanguine view compared to Klein and Majewski (2006), quoted above, who examine the nineteenth century record of private roads.

Like Winston and Shirley (1997), Downs (2004) shows caution:

> Removing major roads from public ownership would by no means remove them from public concern, since they are the backbone of the nation's ground transportation system. As long as the vast majority of American drivers strongly oppose all-lane peak-hour road pricing, American elected officials will never permit its widespread adoption, no matter who owns the roads. (Downs 2004, 162)

C. Opposed to private roads

> Free competition among toll-bridge owners, of the kind necessary to make the conclusion applicable, would require that each bridge be paralleled by an infinite number of others immediately adjacent to it, all the owners being permanently engaged in cutthroat competition. (Hotelling 1938, 260)

> One may get the impression from Professor Knight's article . . . that a private enterprise road system would do the

> job much more efficiently. (I do not know whether Knight would accept this interpretation.) Or one might still have public owned roads, each with a toll gate and a separate manager for each gate; then the state would instruct the managers to maximize their profits. But it is clear that neither of these systems would lead to the ideal allocation. Some frightful state of oligopoly would emerge in both cases; some form of collusion would be likely as the ultimate outcome of the first case. (Walters 1954, 143)

> In our view, publicly provided and efficient priced highways are a better solution than franchised monopolies that tend either to set inefficiently high prices or to sink under the weight of oppressively detailed regulation. (Small, Winston and Evans 1989, 119)

> The problem is that many roads are natural monopolies, and would require regulation that may reduce most of the benefits of private ownership. There is the additional complication that the road network is a network, and investment or charges on one sector will affect the traffic flows and profitability of other sectors. The difficulties of pricing component parts of an integrated network are such that most networks have been retained intact, as with the National Grid, the gas transmission system, British Telecom and, nearer to the present, British RailTrack. (Newbery 1994, 244)[58]

> [T]he welfare gain from managerial efficiency due to private initiatives of road provision via increasingly popular build-operate-transfer (BOT) projects, for instance, should be measured against the welfare loss from monopoly abuse when parallel roads are next to nonexistent. Because many roads possess natural monopoly characteristics and since it is difficult to price various component parts of an integrated road network, the ownership of roads should best reside with the public sector (Newbery, 1994). The market failure resulting from

[58] Instead of privatization Newbery (1994) calls for a public road authority that operates under commercial principles. His arguments are elaborated in Newbery and Santos (1999).

> the common property resource problem where no one really owns the roads would still call for the diligent application of optimal pricing and investment rules by an independent public road authority. Thus commercialization is in order rather than privatization. Vickrey (1996) also insists on 'marketization'—that is, the setting of quasi-market prices which enhances efficiency and acts as signals for (dis)investment—in transport and argues strongly against privatization in transportation. (Hau 1998, 68)

CONCLUSIONS

Economists' model-based analysis of road pricing has developed over a period of decades. Practical proposals were also advanced in the mid-twentieth century, but road pricing continued to remain largely an ivory tower idea. A few operational schemes were launched late in the century, but more attempts failed than succeeded. The recent upsurge of research on road pricing reflects, in large part, belated recognition by economists (and other scholars) that a more practical and pragmatic approach is required for road pricing to take off.

This article has surveyed the literature to determine whether economists support road pricing both in theory and in practice. There is a strong consensus among economists that road pricing ought to be used to reduce and manage congestion. This position has strengthened as technology has reduced the costs of collecting and paying toll charges. Beyond that primary insight, however, there is much disagreement.

The diversity of views described in the survey defies easy summary, but some broad statements are possible.

- Again, there is a strong consensus in favor of using road pricing to manage congestion.
- Most economists now accept short-run marginal cost as the appropriate basis for pricing transportation generally, and roads specifically. However, most economists also entertain departures from SRMC pricing in its strictest sense, and there has been much interest recently in developing practical models of second-best

pricing. There also remains a residual tension between the goal of allocative efficiency, which is best served by SRMC pricing, and the goal of cost recovery which calls for average-cost pricing at some (often ill-defined) level of aggregation over modes and user groups.

- Many economists are concerned about practical aspects of road pricing. The major concerns are difficulties in computing optimal tolls, administration costs and user friendliness. There is less worry about the reliability of tolling technology or the feasibility of differentiating charges over time and space.

- Economists have mixed views on equity. A majority probably hold at least moderately egalitarian moral sentiments. But many oppose the idea of compromising the efficiency of road pricing schemes by granting discounts or exemptions to selected groups, or otherwise distorting prices in an attempt to serve some notion of equity.

- Many economists view road pricing as just one tool that should compete on the merits against other demand-side and supply-side transport policy instruments. The advantages of road pricing as a price instrument in terms of decentralization, information aggregation, etc., must in their view be weighed against its setup and administration costs, and against the proven (if limited) worth of other more direct instruments such as parking bans and pedestrian-only zones as well as fiscal instruments such as vehicle registration fees that do not vary with usage.[59]

- Attitudes towards earmarking of road-user charge revenues are decidedly mixed. Some economists are against earmarking because it reduces budgetary flexibility to respond to changing priorities. Others favour earmarking for roads, for public transport, or for some mix of uses. Still others favour rebating revenues to users, or reducing fuel taxes or fixed charges to maintain revenue neutrality.

[59] This attitude is well-articulated by Arnott (2005, 11): "On one hand, I look forward to seeing what schemes are put in place, and how well they do, and sincerely hope that urban congestion pricing proves to be worth the wait. On the other, I have my doubts that urban congestion pricing will be as effective as most other urban transport economists believe. Whether or not my doubts prove well founded, city tolls are only one element of an effective policy cocktail for dealing with urban traffic congestion. Urban transport economists should broaden their horizons beyond congestion pricing to give due attention to the myriad other congestion-relief policies whose effectiveness can only be improved by the application of sound economics."

- Economists are divided on private roads. Those who write at length about private roads are writing about something that departs greatly from the status quo. They are probably more favourably disposed towards private roads than are economists as a whole, and the sample of authors mentioned here may not be representative partly for this reason.

- Finally, a number of economists support road pricing, but are pessimistic that it will ever become widespread. The doubters include some of the leaders in developing the theory of road pricing in the 1960s.

Several recent favourable developments have boosted interest for road pricing amongst policymakers and researchers, and new road pricing experiments or toll-road proposals are announced frequently. Some economists have caught the enthusiasm. In the preface to their book, Arnott, Rave and Schob (2005) explain how the steadfast support for marginal-cost pricing from the European Union, and success with London's congestion charging scheme, induced them to moderate the largely negative view of road pricing they held a few years ago.

Singapore's electronic road pricing scheme has demonstrated how tolls that are differentiated by vehicle type and time of day can be implemented on a network of roads using relatively sophisticated and user-friendly technology. As Christainsen (2006) notes, "In other countries, roads, medical care, housing, and many other goods and services are systematically underpriced insofar as their prices are controlled by government …. Singapore essentially broke the mold with respect to roadways, and the precise political conditions under which such breakthroughs occur, and endure, deserve more research on the part of political scientists as well as economists. With respect to roadways, one can only speculate whether those conditions are now close to being met in additional cities around the world" (Christainsen 2006, 87). The US Value Pricing Pilot program, and its centrepiece HOT lane facility projects, have also received a largely positive reception. As O'Sullivan (2003) remarks, "These recent experiences with congestion pricing are promising. Travelers respond to higher prices by changing their travel behavior in ways that

decrease traffic volume and improve the efficiency of travel" (O'Sullivan 2003, 267).[60]

The pessimistic assessment by Borins (1988) of Hong Kong's electronic road pricing initiative was noted earlier. However, Mylvaganam and Borins (2004) are more upbeat in their account of Toronto's Highway 407 toll road. They remark that both limited access highways and central cities have been successfully tolled and they "encourage transportation policy makers in Canada . . . to consider both" (131).

This is an exciting time for road pricing. Opponents may hope that present efforts will fail and another chapter in the long history of unsuccessful road pricing attempts will be written. Devotees of Pigou and Vickrey will hope that public toll roads or other forms of road pricing will flourish to tame the congestion beast. And admirers of Friedman and Boorstin will look forward to a resurrection of private toll roads that ruled in the turnpike era.

The potential for better policy is open. J. Michael Thomson suggests that it is up to economists:

> Immense developments have been made in the technology of road pricing, and there is little doubt that a sophisticated and economically efficient system of road pricing could now be introduced as soon as it was required, by a city with a sound administration and good law enforcement …. But there is still an educational gap to be bridged: the gap between economic truth and public comprehension. If it really is true that most, if not all, the people in a city will benefit from a system of road pricing, it is for economists to explain this truth so that politicians, administrators, and members of the public can understand it. Can economists respond to this challenge? (Thomson 1998, 109)

[60] Using data from State Route 91 in California, Small et al. (2006) find that with their base-case parameter values the toll price elasticity of demand to use the toll lanes is roughly one in absolute value. They also demonstrate how accounting for preference heterogeneity among motorists affects the elasticity estimate.

APPENDIX A: ECONLIT SEARCH

Figure 1 in the text was assembled using the ECONLIT database. According to the ECONLIT software the database "covers the world's economic literature. It indexes over 400 major journals as well as articles in collective volumes (essays, proceedings, etc.), books, book reviews, dissertations, and working papers." However, ECONLIT excludes a number of mainly engineering-oriented transportation journals in which articles on road pricing have been published; in particular *Highway Research Record*, *ITE Journal*, *ITS Journal*, *Journal of Transportation Engineering*, *TR News*, *Traffic Engineering and Control*, *Transport Reviews*, *Transportation Quarterly*, *Transportation Research Series C* (Emerging Technologies), *Transportation Research Series F* (Traffic Psychology and Behaviour), *Transportation Research Record* and *Transportation Science*. Figure 1 therefore does not provide a complete count of the literature on road pricing, although it probably captures most of the research by economists.

For the purpose of this survey journal articles, book chapters, comments, rejoinders and conference proceedings were included in the counts. Working papers were excluded unless they appeared in a series maintained by a well-known institution (e.g. the World Bank or the Institute of Economic Affairs) and are not known to have been published elsewhere. Books were counted as one item unless they constituted an edited volume, in which case individual chapters were counted separately. The counts exclude unpublished MA or PhD dissertations, reprints of articles that were published earlier, and editorials except journal editorials for special issues about road pricing.

The search was conducted by title. Alternative sets of keywords were tested with the goal of including known articles with unusual, or frugal, titles without including hordes of irrelevant items. The final choice was: {congestion OR cordon OR (financing AND roads) OR (highway* AND (pricing OR privat*)) OR (marginal AND cost AND pricing) OR (marginal AND cost AND tax*) OR (peak AND load) OR (pricing AND transport*) OR ((road* OR highway*) AND (pric* OR privat* OR charg*)) OR toll*}.

Hits for each year were saved and the titles (and abstracts if available) were then read to assess whether the article concerned road pricing. Any form of direct road user charges was included; i.e. tolls on individual highways or highway lanes, area licenses or charges, cordon tolls, distance-based charges, time-based charges, etc. Charges for any purpose were included; i.e. pricing of congestion, road damage, environmental or other

road vehicle externalities; tolling to raise revenue, and so on. Charges on any type of motorized vehicle were included, but not charges on bicycles, pedestrians or other forms of non-motorized transport. Indirect road charges were excluded; i.e. registration and license fees, fuel taxes, tire taxes, etc. Shadow tolls and tradable driving permits were also excluded, as were parking fines and parking charges. Theoretical studies on peak-load pricing were excluded unless they included road pricing as an extended example.

For about 10 percent of the search hits it was not possible to determine from the title or abstract (if an abstract was provided) whether the criteria described in the previous paragraph were met. In most of these cases a copy of the item was obtained to make the determination. In a few cases the item was not available from the University of Alberta library, and a request for interlibrary loan was either unsuccessful or still outstanding at the end of January, 2006. If so, a judgment call was made on the basis of the information available.

The results of the search are reported in Table 1 below. Over the 36-year period the total number of annual items in the ECONLIT database grew more than seven-fold from 4,525 to 33,724. The average number of hits obtained using the keywords was about 30 per year, and the average number of articles about road pricing was 9 per year. A linear regression against time was performed of the fraction of hits on road pricing (column [5] in Table 1). This yielded a coefficient estimate for time of 0.002 (t-statistic 2.41, P=0.0215). Another linear regression was performed for the number of articles about road pricing as a fraction of all items in the database (column [6] in Table 1).This yielded a coefficient estimate for time of $-2x10^{-6}$ (t-statistic -0.42, P=0.676). Thus, articles about road pricing became more prevalent relative to other articles that were captured by the keywords (in particular, theoretical articles about peak-load pricing or applied articles concerning time-of-day electricity pricing). But, perhaps surprisingly, entries about road pricing did not increase in number as a fraction of all entries over the full 1969-2004 period.

Table 1: ECONLIT Database Entries

[1]	[2]	[3]	[4]	[5]	[6]
Year	Database entries			= [4]/[3]	= [4]/[2]
	Total	**Keyword**	**Road pricing**	**Road pricing as fraction of keywords**	**Road pricing as frac. of total [10^{-4}]**
1969	4,525	14	5	0.36	11.0
1970	5,098	16	2	0.13	3.9
1971	5,101	17	4	0.24	7.8
1972	5,821	7	1	0.14	1.7
1973	6,147	13	3	0.23	4.9
1974	6,184	8	1	0.13	1.6
1975	6,264	14	6	0.43	9.6
1976	6,669	18	5	0.28	7.5
1977	7,332	22	7	0.32	9.5
1978	7,861	17	4	0.24	5.1
1979	8,123	18	0	0.00	0.0
1980	8,662	17	5	0.29	5.8
1981	8,959	23	7	0.30	7.8
1982	8,982	17	6	0.35	6.7
1983	9,916	24	2	0.08	2.0
1984	14,896	23	3	0.13	2.0
1985	15,345	24	1	0.04	0.7
1986	17,475	20	4	0.20	2.3
1987	18,481	18	7	0.39	3.8
1988	21,988	27	4	0.15	1.8
1989	21,920	22	3	0.14	1.4
1990	23,496	33	6	0.18	2.6
1991	26,200	27	5	0.19	1.9
1992	27,407	24	8	0.33	2.9
1993	27,504	27	8	0.30	2.9
1994	28,670	28	3	0.11	1.0
1995	31,126	43	22	0.51	7.1
1996	33,261	58	13	0.22	3.9
1997	34,609	45	8	0.18	2.3
1998	35,385	62	20	0.32	5.7
1999	38,307	59	25	0.42	6.5
2000	39,651	48	25	0.52	6.3
2001	37,947	68	26	0.38	6.9
2002	38,125	75	30	0.40	7.9
2003	39,276	62	16	0.26	4.1
2004	33,724	54	29	0.54	8.6
Mean		**30.3**	**9.0**	**0.26**	**4.7**

Sources: Author's compilation

APPENDIX B: COVERAGE OF ROAD PRICING IN TRANSPORTATION ECONOMICS TEXTBOOKS

Thomson (1998, 93) notes that early transportation textbooks (1941, 1946, 1958) did not even mention urban traffic congestion, let alone road pricing. Table 2 summarizes the coverage and assessments of nine more recent transport economics textbooks.[61]

Table 2: Coverage of road pricing in selected transportation economics textbooks (chronological order)

Textbook	Coverage of road pricing	Attitudes
Gwilliam and Mackie (1975)	Chaps. 6&10. 3-4 pp.	Agnostic. Road pricing should be assessed against a range of policy instruments.
Mohring (1976)	Chaps. 3&6 ~7 pp.	Does not explicitly advocate tolls. Implicitly identifies their benefits.
Frankena (1979)	Chap. 4. 14 pp.	Relatively favourable.
Glaister (1981)	Chap. 5. 5-6 pp. Considerable additional space on peak-load pricing and 2nd-best pricing of public transport with unpriced auto. congestion.	Relatively favourable. Advocates tolling. Against private roads in absence of congestion.
Small (1992)	Chap. 4. ~18 pp. Devotes some of this space to comparison of results from different models.	Favourable towards road pricing, although no explicit statement to this effect.

[61] Road pricing is also covered in other economics textbooks; notably those about urban economics.

Small (1992) (Cont.)	Private highways: 3 pp.	Private highways: reservations about optimality of pricing and investment strategies of a profit-maximizing highway owner with market power
Boyer (1998)	Chap. 10. 5 pp.	Enthusiastic. Considers some form of congestion pricing in the future as inevitable
McCarthy (2001)	Chaps. 10&11: ~45 pp. Nothing on private roads	Reviews selected academic studies in detail. Shows approval of pricing highway congestion and infrastructure damage.
Quinet and Vickerman (2004)	Chaps. 7&10. 5 pp.	Moderately favourable.
Small and Verhoef (2006) (Revised edition of Small (1992), in process.)	Chap. 4. 60 manuscript pp. Private ownership of highways: Covered in Chap. 6. Draft incomplete	First-best pricing, second-best pricing and congestion pricing in practice. Generally favourable.

Source: Author's compilation

Except for McCarthy (2001), who adopts a case-study approach by summarizing in detail several articles about road pricing, there is no obvious trend over the 30-year span between Gwilliam and Mackie (1975) and Quinet and Vickerman (2004) in either attitudes towards road pricing or the amount of space devoted to the subject. Small and Verhoef (2006) stand out in providing a much more comprehensive and detailed treatment of road pricing—mainly in the role of pricing congestion. Chapter 5 on

investment devotes further attention to pricing in the context of the link between pricing and capacity choice decisions. Chapter 6, which covers industrial organization in urban transportation, includes private ownership of highways. Chapter 6 was incomplete when the final version of this review was prepared.

REFERENCES

Abouchar, Alan. 1987. How do We Price America's Highways? *Challenge* 30(2): 52-54.

Arnott, Richard A. 1998. William S. Vickrey. *International Tax and Public Finance* 5(1): 93-113.

Arnott, Richard A. 2005. City Tolls**—**One Element of an Effective Policy Cocktail. *Dice Report: Journal of Institutional Comparisons* 3(3): 5-11. Link. (cited: November 20, 2005).

Arnott, Richard and Marvin Kraus. 1998. When Are Anonymous Congestion Charges Consistent with Marginal Cost Pricing? *Journal of Public Economics* 67: 45-64.

Arnott, Richard A., Tilmann Rave and Ronnie Schöb. 2005. *Alleviating Urban Traffic Congestion*. Cambridge: MIT Press.

Banister, David. 2002. The Integration of Road Pricing with Land Use Planning. Paper prepared for second seminar of the IMPRINT-EUROPE Thematic Network: Implementing Reform on Transport Pricing: Identifying Mode-Specific Issues. Brussels, May, 14-15.

Baumol, William J. and Dietrich Fischer. 1987. Peak Pricing, Congestion, and Fairness. In *Arrow and the Foundations of the Theory of Economic Policy*. New York: New York University Press, 382-409.

Beckmann, Martin, C. Bartlett McGuire and Christopher B. Winsten. 1956. *Studies in the Economics of Transportation*. New Haven, CT: Yale University Press.

Beesley, Michael E. 1973. Road Pricing: Economics, Techniques and Policy. In *Urban Transport: Studies in Economic Policy*. London: Butterworths, 223-286.

Beesley, Michael E. and David Hensher. 1990. Private Tollroads in Urban Areas: Some Thoughts on the Economic and Financial Issues. *Transportation* 16(4): 329-341.

Beesley, Michael E. and Gabriel J. Roth. 1962-63. Restraint of Traffic in Congested Areas. *The Town Planning Review* 33: 184-196.

Berg, John. T. 2003. Listening to the Public: Assessing Public Opinion about Value Pricing. State and Local Government Program, Hubert H. Humphrey Institute of Public Affairs.

Blaug, Mark. 1985. Marginal Cost Pricing: No Empty Box. In *Public Choice, Public Finance and Public Policy: Essays in Honour of Alan Peacock*. Oxford and New York: Blackwell, 15-30.

Block, Walter. 1979. Free Market Transportation: Denationalizing the Roads. *The Journal of Libertarian Studies* 3(2): 209-238.

Block, Walter. 1980. Congestion and Road Pricing. *The Journal of Libertarian Studies* 4(3): 299-330.

Block, Walter. 1996. Road Socialism. *International Journal of Value-Based Management* 9: 195-207.

Borins, Sanford F. 1988. Electronic Road Pricing: An Idea Whose Time May Never Come. *Transportation Research A* 22A: 37-44.

Boyce, David, Hani S. Mahmassani and Anna Nagurney. 2005. A Retrospective on Beckmann, McGuire and Winsten's Studies in the Economics of Transportation. *Papers in Regional Science* 84(1): 85-103.

Boyer, Ken D. 1998. *Principles of Transportation Economics*. Reading, Mass: Addison Wesley Longman.

Braybrooke, David. 1974. *Traffic Congestion Goes Through the Issue-Machine; A Case Study in Issue Processing, Illustrating a New Approach*. London: Routledge & Kegan Paul.

Brittan, Samuel. 1973. *Is There an Economic Consensus?* London: Macmillan.

Buchanan, James M. 1952. The Pricing of Highway Services. *National Tax Journal*. V(2): 97-106.

Buchanan, James M. 1956. Private Ownership and Common Usage: The Road Case Re-examined. *Southern Economic Journal*. 22: 305-316.

Buchanan, James M. 1996. Foreword. In *Roads in a Market Economy*. Gabriel J. Roth. Aldershot, Hants: Avebury Technical, xv-xvi.

Button, Kenneth J. 2004. The Rationale for Road Pricing. In *Road Pricing: Theory and Evidence, Research in Transportation Economics 9*, ed. Georgina Santos. Amsterdam: Elsevier Science, 3-25.

Button, Kenneth J. 2004. Road Pricing: A Research Project Report for the National ITS Implementation Research Center. Center for Transportation Policy, Operations and Logistics, School of Public Policy, George Mason University. March.

Button, Kenneth J. 2006. How Should the Revenues from Congestion Pricing be Spent?. In *Street Smart: Competition, Entrepreneurship and the Future of Roads*, ed. Gabriel J. Roth. The Independent Institute, Transaction Publishers, New Brunswick, US, and London, UK., 225-242.

Cadin, Michelle S. and Walter Block. 1997. Privatize Public Highways. *The Freeman* 47(2): 96-97.

Calfee, John and Clifford Winston. 1998. The Value of Automobile Travel Time: Implications for Congestion Policy. *Journal of Public Economics* 69: 83-102.

Cameron, Michael W. 1994. Efficiency and Fairness on the Road: Strategies for Unsnarling Traffic in Southern California. Oakland: Environmental Defense Fund.

Chernick, H. and A. Reschovsky. 1997. Who Pays the Gasoline Tax? *National Tax Journal* 50(2): 233-259.

Christainsen, Gregory B. 2006. Road Pricing in Singapore after 30 Years. *Cato Journal* 26(1): 71-88.

Clark, J. Maurice. 1923. *Studies in the Economics of Overhead Costs*. Chicago & London: University of Chicago Press.

Clemens, Eli Winston. 1941. Price Discrimination in Decreasing Cost Industries. *American Economic Review* 31(4): 794-802.

Coase, Ronald H. 1946. The Marginal Cost Controversy. *Economica*, n.s. 13, August, 169-182. Reprinted as Chapter 4 (75-93) in Coase's *The*

Firm, the Market, and the Law. Chicago, University of Chicago Press, 1988.

Colgan Charles S. and Gary Quinlin. 1997. The Catch-22 of Congestion Pricing. *Transportation Quarterly* 51(4): 117-133.

Dafermos, Stella C. and Frederick T. Sparrow. 1971. Optimal Resource Allocation and Toll Patterns in User-Optimised Transport Networks. *Journal of Transport Economics and Policy* 5(2): 184-200.

Day, Alan. 1998. The Case for Road Pricing. *Institute for Economic Affairs*, December: 5-8.

De Vany, Arthur and Thomas R. Saving. 1980. Competition and Highway Pricing for Stochastic Traffic. *Journal of Business* 53(1): 45-60.

Decorla-Souza, Patrick. 2004. Recent U.S. Experience: Pilot Projects. In *Road Pricing: Theory and Evidence*, *Research in Transportation Economics* 9, ed. Georgina Santos. Amsterdam: Elsevier Science, 283-308.

Demsetz, Harold. 1969. Information and Efficiency: Another Viewpoint. *Journal of Law and Economics* 12(1): 1-22.

Derycke, Pierre-Henri. 1998. Road Pricing—Analysis and Policies (A Historical Perspective). *Recherches Economiques de Louvain* 64(1): 63-74.

Downs, Anthony. 1962. The Law of Peak-Hour Expressway Congestion. *Traffic Quarterly* 16: 393-409.

Downs, Anthony. 2004. *Still Stuck in Traffic: Coping with Peak-Hour Traffic Congestion*. Washington, D.C.: Brookings Institution Press.

Drèze, Jacques H. 1994. Introduction to Part III, Marginal Cost Pricing. In *Public Economics: Selected Papers by William Vickrey*, ed. Richard Arnott, Kenneth Arrow, Anthony Atkinson and Jacques H. Drèze, Cambridge: Cambridge University Press, 189-195.

Dupuit, Jules. 1844. De la Mesure de l'utilité des Travaux Publics. *Annales des Ponts et Chaussées: Mémoires et Documents*, 2nd series 8(2): 332-375. Trans. R.M. Barback, "On the Measurement of Utility for Public Works". International Economic Papers, 1952, 2: 83-110.

Dupuit, Jules. 1849. Des Péages. *Annales des Ponts et Chaussées: Mémoires et Documents*, 2nd ser., 17(1), 207-248. Trans. Elizabeth Henderson, On Tolls and Transport Charges. *International Economic Papers*, 1962, 11: 7-31.

Economist. 1989. Make Them Pay. *The Economist*, February 18, 11-12.

Edelson, Noel M. 1971. Congestion Toll under Monopoly. *American Economic Review* 61: 873-882.

Ekelund Jr., Robert E. and Robert F. Hébert. 1999. *Secret Origins of Modern Microeconomics: Dupuit and the Engineers*. Chicago and London: University of Chicago Press.

Elliott, Ward. 1986. Fumbling Toward the Edge of History: California's Quest for a Road-Pricing Experiment. *Transportation Research A* 20A(2): 151-156.

Elliott, Ward. 2000. Greenbacks Über Gridlock: REACH Task Force Shows L.A. How to Save Billions in Smog and Congestion Costs. *Planning and Markets* 3(1). Link. (cited: April 14, 2006).

Else, Peter K. 1986. No Entry for Congestion Taxes? *Transportation Research A* 20A(2): 99-108.

European Commission. 1995. Towards Fair and Efficient Pricing in Transport - Policy Options for Internalising the External Costs of Transport in the European Union. Green Paper, COM(95) 691 final. Brussels: European Commission.

European Commission. 1998. Fair Payment for Infrastructure Use: A Phased Approach to a Common Transport Infrastructure Charging Framework in the EU. White paper, COM(98) 466 final. Brussels: European Commission.

European Commission. 2001. White Paper on European Transport Policy for 2010: Time to Decide. Luxembourg: Office for Official Publications of the European Communities.

Evans, Andrew W. 1992. Road Congestion Pricing: When is it a Good Policy?. *Journal of Transport Economics and Policy* 26, September: 213-244.

Farrell, Séona and Wafaa Saleh. 2005. Road-user Charging and the Modelling of Revenue Allocation. *Transport Policy* 12: 431-442.

Feldstein, Martin. 1972. Distributional Equity and the Optimal Structure of Public Prices. *American Economic Review* 62(1): 32-36.

Fielding, Gordon J. and Daniel B. Klein. 1993. How to Franchise Highways. *Journal of Transport Economics and Policy* 17, May: 113-130.

Foldvary, Fred. 2006. Streets as Private-Sector Public Goods. In *Street Smart: Competition, Entrepreneurship and the Future of Roads*, ed. Gabriel J. Roth. The Independent Institute, New Brunswick, US, and London, UK.: Transaction Publishers, 305-325.

Foster, Christopher. 1963. *The Transport Problem*. London: Croom Helm.

Foster, Christopher. 1974. The Regressiveness of Road Pricing. *International Journal of Transport Economics* 1: 133-141.

Foster, Christopher. 1975. A Note on the Distributional Effects of Road Pricing: A Comment. *Journal of Transport Economics and Policy* 9: 186-187.

Frankena, Mark W. 1979. *Urban Transportation Economics: Theory and Canadian Policy*. Toronto: Butterworths.

Friedlaender, Ann F. 1969. *The Dilemma of Freight Transport Regulation*. Washington, DC: Brookings Institution.

Friedman, David. 1989. *The Machinery of Freedom*, 2nd edition. La Salle: Open Court. [1st edition 1969.]

Friedman, Milton and Daniel Boorstin. 1951. How to Plan and Pay for the Safe and Adequate Highways We Need. Epilogue in Gabriel J. Roth (1996). *Roads in a Market Economy*. Aldershot, Hants: Avebury Technical, 223-245.

Gillen, David. 1997. Efficient Use and Provision of Transportation Infrastructure with Imperfect Pricing: Second Best Rules. In *The Full Costs and Benefits of Transportation*, ed. David L. Greene, Donald W. Jones and Mark A. Delucchi. Berlin: Springer-Verlag, 193-218.

Giuliano, Genevieve. 1992. An Assessment of the Political Acceptability of Congestion Pricing. *Transportation* 19: 335-358.

Giuliano, Genevieve. 1994. Equity and Fairness Considerations of Congestion Pricing. Curbing Gridlock, Peak-Period Fees to Relieve Traffic Congestion, Transportation Research Board Special Report 242: 250-279.

Glaister, Stephen. 1981. *Fundamentals of Transport Economics*. New York: St. Martin's Press.

Glazer, Amihai. 1981. Congestion Tolls and Consumer Welfare. *Public Finance* 36(1): 77-83.

Gómez-Ibáñez, Jose A. 1992. The Political Economy of Highway Tolls and Congestion Pricing. *Transportation Quarterly* 46: 343-360.

Gómez-Ibáñez, Jose A. and John R. Meyer. 1993. *Going Private: The International Experience with Transport Privatization*. Washington, D.C.: The Brookings Institution.

Goodwin, Philip B. 1989. The 'Rule of Three': A Possible Solution to the Political Problem of Competing Objectives for Road Pricing. *Traffic Engineering + Control* 30(10): 495-497.

Goodwin, Philip B. 1997. Solving Congestion. Inaugural Lecture of the Professorship of Transport Policy, University College London. Link. (cited: December 7, 1999).

Gronau, Reuben. 1994. Fuel Taxes and Road-User Charges in LDCs: Some Lessons from Sub-Saharan Africa. *Journal of Transport Economics and Policy* 28(3): 255-273.

Grubbs, Clifton M. 1964. Theory of Spillover Cost Pricing. *Traffic Congestion as a Factor in Road Users Taxation, Highway Research Record* 47: 15-22.

Gulipalli, P.K., Sukumar Kalmanje and Kara M. Kockelman. 2005. Credit-based Congestion Pricing: Expert Expectations and Guidelines for Application. *84th Annual Meeting of the Transportation Research Board*, Paper No. 05-1848, Washington, D.C.

Gwilliam, Kenneth and Peter J. Mackie. 1975. *Economics and Transport Policy*. London: George Allen & Unwin Ltd..

Harford, Timothy. 2006. *The Undercover Economist*. Oxford, New York: Oxford University Press.

Harrington, Winston, Alan. J. Krupnick and Anna Alberini. 2001. Overcoming Public Aversion to Congestion Pricing. *Transportation Research A* 35(2): 87-105.

Hau, Timothy D. 1992. Economic Fundamentals of Road Pricing: A Diagrammatic Analysis. WPS 1070, World Bank Infrastructure and Development Department.

Hau, Timothy D. 1998. Congestion Pricing and Road Investment. In *Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social*

Feasibility, ed. Kenneth J. Button and Erik T. Verhoef. Cheltenham, UK: Edward Elgar, 39-78.

Hau, Timothy D. 2005a. Economic Fundamentals of Road Pricing: A Diagrammatic Analysis, Part I—Fundamentals. *Transportmetrica* 1(2): 81-115.

Hau, Timothy D. 2005b. Economic Fundamentals of Road Pricing: A Diagrammatic Analysis, Part II—Relaxation of Assumptions. *Transportmetrica* 1(2): 119-149.

Hensher, David A. 2003. Congestion Charging: What Sydney can Learn from London. On Line Opinion. Link. (cited: November 17, 2004).

Hensher, David A. and Sean M. Puckett. 2005. Road User Charging: The Global Relevance of Recent Developments in the United Kingdom. *Transport Policy* 12: 377-383.

Higgins, Thomas. 1986. Road Pricing Attempts in the United States. *Transportation Research A* 20A(2): 145-150.

Hotelling, Harold. 1938. The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates. *Econometrica* 6(3): 242-269.

Hutt, William Harold. 1971. *Politically Impossible . . .? An Essay on the Supposed Electoral Obstacles Impeding the Translation of Economic Analysis into Policy, or, Why Politicians Do Not Take Economic Advice*. London: Institute of Economic Affairs.

Ison, Stephen. 2004. *Road User Charging: Issues and Policies*. Hampshire, UK: Ashgate Publishing Ltd.

Johansson, Lars-Olof; et al. 2003. Goal Conflicts in Political Decisionmaking: A Survey of Municipality Politicians' Views of Road Pricing. *Environment and Planning C* 21(4): 615-624.

Keeler, Theodore E. and Kenneth A. Small. 1977. Optimal Peak-Load Pricing, Investment, and Service Levels on Urban Expressways. *Journal of Political Economy* 85(1): 1-25.

King, David A., Michael Manville and Donald Shoup. 2006. Political Calculus of Congestion Pricing. 85th Annual Meeting of the Transportation Research Board, Washington, D.C. Conference CD Paper No. 06-2703.

Klein, Daniel B. 1998. Planning and the Two Coordinations, With Illustration in Urban Transit. *Planning and Markets* 1(1). Link. (cited: February 22, 2006).

Klein, Daniel B. and John Majewski. 2006. America's Toll Roads Heritage: The Achievements of Private Initiative in the Nineteenth Century. In *Street Smart: Competition, Entrepreneurship and the Future of Roads*, ed. Gabriel J. Roth. The Independent Institute, New Brunswick, US, and London, UK., Transaction Publishers, 277-303.

Knight, Frank. 1924. Some Fallacies in the Interpretation of Social Costs. *Quarterly Journal of Economics* 38(4): 582-606.

Kockelman, Kara and Sukumar Kalmanje. 2005. Credit-based Congestion Pricing: A Proposed Policy and the Public's Response. *Transportation Research A* 39A: 671-690.

Lave, Charles A. 1995. The Demand Curve under Road Pricing and the Problem of Political Feasibility: Author's Reply. *Transportation Research A* 29A(6): 464-465.

Layard, Richard. 1977. The Distributional Effects of Congestion Taxes. *Economica* 44: 297-304.

Lévy-Lambert, H. 1968. Tarification des Services à Qualité Variable: Application aux Péages de Circulation. *Econometrica* 36: 564-574.

Lindsey, Robin and Erik T. Verhoef. 2001. Traffic Congestion and Congestion Pricing. In *Handbook of Transport Systems and Traffic Control*, ed. Kenneth J. Button and David A. Hensher. Oxford: Elsevier Science, 77-105.

Littlechild, Steven C. 1973. Myopic Investment Rules and Toll Charges in a Transport Network. *Journal of Transport Economics and Policy* 7, May: 194-204.

Marchand, Maurice G. 1968. A Note on Optimal Tolls in an Imperfect Environment. *Econometrica* 36(3-4): 575-581.

McCarthy, Patrick S. 2001. *Transportation Economics: Theory and Practice - A Case Study Approach*. Malden, MA: Blackwell Publishers.

McDonald, John F., Edmond L. d'Ouville and L. Nan Liu. 1999. *Economics of Urban Highway Congestion and Pricing*. Boston: Kluwer Academic Publishers.

McQuaid, Ronald and Margaret Grieco. 2005. Edinburgh and the Politics of Congestion Charging: Negotiating Road User Charging with Affected Publics. *Transport Policy* 12: 475-476.

Meiburg, Charles O. 1963. An Economic Analysis of Highway Services. *Quarterly Journal of Economics* 77: 648-656.

Meyer, John R., John F. Kain and Martin Wohl. 1965. *The Urban Transportation Problem*. Cambridge, Mass.: Harvard University Press.

Meyer, John R., Merton J. Peck, John Stenason and Charles Zwick. 1959. *The Economics of Competition in the Transportation Industries*. Cambridge: Harvard University Press.

Mills, David E. 1981. Ownership Arrangements and Congestion-Prone Facilities. *American Economic Review* 71(3): 493-502.

Mills, Edwin S. 1998. Excess Commuting in U.S. Metropolitan Areas. In *Network Infrastructure and the Urban Environment*, ed. Lars Lundqvist, Lars-Goran Mattsson and Tschangho J. Kim. Berlin: Springer-Verlag, 72-83.

Mohring, Herbert. 1964. Relation Between Optimum Congestion Tolls and Present Highway User Charges. *Traffic Congestion as a Factor in Road-User Taxation, Highway Research Record* 47: 1-14.

Mohring, Herbert. 1976. *Transportation Economics*. Ballinger, Cambridge, Mass.

Mohring, Herbert. 1991. Letter to Anthony Downs. In *Still Stuck in Traffic: Coping with Peak-Hour Traffic Congestion*, Anthony Downs. Washington, D.C.: Brookings Institution Press (2004). 422, footnote 14.

Mohring, Herbert. 1999. Congestion. In *Essays in Transportation Economics and Policy: A Handbook in Honor of John R. Meyer*, ed. Jose A. Gómez-Ibáñez, William B. Tye and Clifford Winston. Washington, D.C.: Brookings Institution Press, 181-221.

Mohring, Herbert and David Anderson. 1994. Congestion Pricing for the Twin Cities Metropolitan Area. Working paper, Dept. of Economics, U. of Minnesota.

Mohring, Herbert and Mitchell Harwitz. 1962. *Highway Benefits: An Analytical Framework*. Evanston, Il: Northwestern University Press.

Mylvaganam, Chandran and Sanford Borins. 2004. *If you Build it ... Business, Government and Ontario's Electronic Toll Highway*. Toronto:

University of Toronto Press, University of Toronto Centre for Public Management.

Nash, Chris. 2003. Marginal Cost and Other Pricing Principles for User Charging in Transport: A Comment. *Transport Policy* 10(4): 345-348.

Nelson, James C. 1962. The Pricing of Highway, Waterway, and Airway Facilities. *American Economic Review (Papers and Proceedings)* 52(2): 426-432.

Netzer, Richard. 1952. Toll Roads and the Crisis in Highway Finance. *National Tax Journal* V(2): 107-119.

Newbery, David. 1988. Road Damage Externalities and Road User Charges. *Econometrica* 56(2): 295-316.

Newbery, David. 1989. Cost Recovery from Optimally Designed Roads. *Economica* 56: 165-185.

Newbery, David. 1990. Pricing and Congestion: Economic Principles Relevant to Pricing Roads. *Oxford Review of Economic Policy* 6(2): 22-38.

Newbery, David. 1994. The Case for a Public Road Authority. *Journal of Transport Economics and Policy* 28(3): 235-253.

Newbery, David and Georgina Santos. 1999. Road Taxes, Road User Charges and Earmarking. *Fiscal Studies* 20(2): 103-132.

O'Sullivan, Arthur M. 2003. *Urban Economics*. Boston: McGraw-Hill Irwin.

Oldridge, Brian. 1995. Congestion Metering in Cambridge City, United Kingdom. In *Road Pricing: Theory, Empirical Assessment and Policy*, ed. Börje Johansson and Lars-Göran Mattsson. Boston: Kluwer, 131-140.

Paine, C.L. 1937. Some Aspects of Discrimination by Public Utilities. *Economica* n.s. 4(16): 425-439.

Peterson, Shorey. 1932. Highway Policy on a Commercial Basis. *Quarterly Journal of Economics* 46(3): 417-443.

Peterson, Shorey. 1950. The Highway from the Point of View of the Economist. In *Highways in Our National Life*, ed. Jean Labatut and Wheaton Joshua Lane. Princeton: Princeton University Press, 190-200.

Philbrook, Clarence. 1953. 'Realism' in Policy Espousal. *American Economic Review* 43(5): 846-859.

Pigou, Arthur C. 1912. *Wealth and Welfare*. London: Macmillan.

Pigou, Arthur C. 1920. *The Economics of Welfare*. London: Macmillan.

Poole, Robert. W., Jr. and C. Kenneth Orski. 2003. HOT Networks: A New Plan for Congestion Relief and Better Transit. Reason Public Policy Institute Policy Study No. 305.

Poterba, James M. 1991. Is the Gasoline Tax Regressive? In *Tax Policy and the Economy*, Vol. 5, ed. David Bradford. Cambridge, MA: MIT Press, 145-164.

Quinet, Emile and Roger Vickerman. 2004. *Principles of Transport Economics*. Cheltenham and Northampton, Mass: Edward Elgar.

Ramsey, Frank. 1927. A Contribution to the Theory of Taxation. *Economic Journal* 37: 47-61.

Richards, Martin G. 2006. *Congestion Charging in London: The Policy and the Politics*. Houndsmill, Basingstoke, Hampshire and New York: Palgrave Macmillan.

Richardson, Harry W. 1974. A Note on the Distributional Effects of Road Pricing. *Journal of Transport Economics and Policy* 8: 82-85.

Richardson, Harry W. 1975. A Note on the Distributional Effects of Road Pricing: A Rejoinder. *Journal of Transport Economics and Policy* 9: 188.

Richardson, Harry W. and Chang-Hee Christine Bae. 1998. The Equity Impacts of Road Congestion Pricing. In *Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility*, ed. Kenneth J. Button and Erik T. Verhoef. Cheltenham, UK: Edward Elgar, 247-262.

Roth, Gabriel. 1966. *A Self-Financing Road System*. London, England: The Institute of Economic Affairs.

Roth, Gabriel. 1967. *Paying for Roads: The Economics of Traffic Congestion*. Middlesex, England: Penguin.

Roth, Gabriel. 1996. *Roads in a Market Economy*. Aldershot, Hants: Avebury Technical.

Roth, Gabriel. 1998. Road Pricing in a Free Society. *Institute for Economic Affairs* December: 9-14.

Roth, Gabriel, ed. 2006. *Street Smart: Competition, Entrepreneurship and the Future of Roads*. The Independent Institute, New Brunswick, US, and London, UK: Transaction Publishers.

Roth, Gabriel. 2006. Why Involve the Private Sector in the Provision of Public Roads? In *Street Smart: Competition, Entrepreneurship and the Future of Roads*, ed. Gabriel J. Roth. The Independent Institute, New Brunswick, US, and London, UK., Transaction Publishers, 3-21.

Rothbard, Murray. 1973. *For a New Liberty*. New York: Macmillan.

Rothengatter, Werner. 2003. How Good is First Best? Marginal Cost and Other Pricing Principles for User Charging in Transport. *Transport Policy* 10(2): 121-130.

Sandholm, William H. 2002. Evolutionary Implementation and Congestion Pricing. *Review of Economic Studies* 69(3): 667-689.

Sandholm, William H. 2005. Negative Externalities and Evolutionary Implementation. *Review of Economic Studies* 72(3): 885-915.

Sandholm, William H. 2006. Pigouvian Pricing and Stochastic Evolutionary Implementation. *Journal of Economic Theory*, forthcoming.

Santos, Georgina, ed. 2004 *Road Pricing: Theory and Evidence*, *Research in Transportation Economics 9*, Amsterdam: Elsevier Science.

Schade, Jens and Bernhard Schlag, eds. 2003. *Acceptability of Transport Pricing Strategies*. Amsterdam: Elsevier.

Sharp, Clifford. 1966. Congestion and Welfare—An Examination of the Case for a Congestion Tax. *Economic Journal* 76: 806-817.

Shoup, Donald C. 2005. *The High Cost of Free Parking*. Chicago, Illinois and Washington, D.C.: APA Planners Press.

Singapore Land Transport Authority. Link. (cited: February 11, 2006).

Small, Kenneth A. 1992. *Urban Transportation Economics*. In *Fundamentals of Pure and Applied Economics*, ed. Jacques Lesourne and Hugo Sonnenschein. Chur, Switzerland: Harwood Academic Publishers.

Small, Kenneth A. and Xuehao Chu. 2003. Hypercongestion. *Journal of Transport Economics and Policy* 37(3): 319-352.

Small, Kenneth A. and Jose A. Gómez-Ibáñez. 1998. Road Pricing for Congestion Management: The Transition from Theory to Policy. In

Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility, ed. Kenneth J. Button and Erik T. Verhoef. Cheltenham, UK: Edward Elgar, 213-246.

Small, Kenneth A. and Erik T. Verhoef. 2006. *Urban Transportation Economics*, second edition. Chapters 2 and 3. Link.

Small, Kenneth A. and Clifford Winston. 1988. Optimal Highway Durability. *American Economic Review* 78(3): 560-569.

Small, Kenneth A., Clifford Winston and Carol A. Evans. 1989. *Road Work*. Washington D.C.: Brookings.

Small, Kenneth A., Clifford Winston and Jia Yan. 2006. Differentiated Road Pricing, Express Lanes, and Carpools: Exploiting Heterogeneous Preferences in Policy Design. Forthcoming, *Brookings-Wharton Papers on Urban Affairs*.

Smith, Adam. 1937. *An Inquiry into the Nature and Causes of the Wealth of Nations*, Modern Library edition. Edwin Cannan ed. New York: The Modern Library.

Smith, Roger S. 1975. Highway Pricing and Motor Vehicle Taxation in Developing Countries: Theory and Practice. *Finanzarchiv* 33(3): 451-474.

St. Clair, Guilford 1964. Congestion Tolls - An Engineer's Viewpoint. *Traffic Congestion as a Factor on Road Users Taxation, Highway Research Record* 47: 66-111.

Thomson, J. Michael. 1998. Reflections on the Economics of Traffic Congestion. *Journal of Transport Economics and Policy* 32(1): 93-112.

Toll Roads News. Link. (cited: February 11, 2006).

Transportation Research Board. 2003. Road Value Pricing: Traveler Response to Transport System Changes. *Transit Cooperative Research Program* Report 95, Transportation Research Board. Link. (cited: November 20, 2005).

Turvey, Ralph. 1968. Peak-load Pricing. *Journal of Political Economy* 76(1): 101-113.

U.K. Ministry of Transport. 1964. Road Pricing: The Economic and Technical Possibilities. London: H.M. Stationery Office.

Value Pricing Pilot Program. 2005. Link. (cited: November 20, 2005).

Verhoef, Erik T. 2002a. Second-best Congestion Pricing in General Networks: Heuristic Algorithms for Finding Second-Best Optimal Toll Levels and Toll Points. *Transportation Research B* 36B(8): 707-729.

Verhoef, Erik T. 2002b. Second-best Congestion Pricing in General Static Transportation Networks with Elastic Demands. *Regional Science and Urban Economics* 32: 281-310.

Verhoef, Erik T., Peter Nijkamp and Piet Rietveld. 1997. The Social Feasibility of Road Pricing: A Case Study for the Randstad Area. *Journal of Transport Economics and Policy* 31(3): 255-276.

Verhoef, Erik T. and Kenneth A. Small. 2004. Product Differentiation on Roads: Constrained Congestion Pricing with Heterogeneous Users. *Journal of Transport Economics and Policy* 38(1): 127-156.

Vickrey, William S. 1948. Some Objections to Marginal-Cost Pricing. *Journal of Political Economy* 56: 218-238.

Vickrey, William S. 1963. Pricing in Urban and Suburban Transport. *American Economic Review* 53: 452-465.

Vickrey, William S. 1967. Optimization of Traffic and Facilities. *Journal of Transport Economics and Policy* 1: 123-136.

Vickrey, William S. 1968. Congestion Charges and Welfare: Some Answers to Sharp's Doubts. *Journal of Transport Economics and Policy* 2: 107-118.

Vickrey, William S. 1971. Responsive Pricing of Public Utility Services. *Bell Journal of Economics and Management Science* 2: 337-346.

Vickrey, William S. 1987. Marginal and Average Cost Pricing. *The New Palgrave: A Dictionary of Economics Vol. 1*. New York: Macmillan, 311-318.

Vickrey, William S. 1993. Point of View: Principles and Applications of Congestion Pricing. *TR News* 167: 4-5.

Vickrey, William S. 1996. Privatization and Marketization of Transportation. In *Privatizing Transportation Systems*, ed. Simon Hakim, Paul Seidenstat and Gary W. Bowman. Westport, Connecticut: Praeger, 221-248.

Walters, Alan A. 1954. Track Costs and Motor Taxation. *Journal of Industrial Economics* 2(2): 135-146.

Walters, Alan A. 1961. The Theory and Measurement of Private and Social Cost of Highway Congestion. *Econometrica* 29: 676-699.

Walters, Alan A. 1968. *The Economics of Road User Charges*. World Bank Occasional Paper Number 5, International Bank for Reconstruction and Development. Baltimore, Maryland: Johns Hopkins University Press.

Walters, Alan A. 1988. A Life Philosophy. *British Library of Political and Economic Science*, June 1988: 18-24. URL not recorded. (cited: October 22, 2004). No longer available on Internet.

Walters, Alan A. 2002. Petrol Taxes and Congestion Charges. *Economic Affairs* 22: 40.

Winston, Clifford and Chad Shirley. 1998. *Alternate Route: Towards Efficient Urban Transportation*. Washington, D.C.: Brookings Institution.

Yang, Hai and Hai-Jun Huang. 2005. *Mathematical and Economic Theory of Road Pricing*. Amsterdam: Elsevier.

Zettel, Richard M. and Richard R. Carll. 1964. The Basic Theory of Efficiency Tolls: The Tolled, the Tolled-off, and the Un-tolled. *Traffic Congestion as a Factor in Road-User Taxation, Highway Research Record* 47: 46-65.

ABOUT THE AUTHOR

Robin Lindsey is a professor of economics at the University of Alberta where he has worked since 1982. He received his PhD in economics from Princeton. His academic specialty is transportation economics, and his current research interests include traffic congestion, road pricing and private toll roads. He has also published research on pricing of parking, advanced traveller information systems, retail market competition, predatory pricing and the economics of exhaustible resources. He has published in *American Economic Review*, *Economica*, *European Economic Review*, *International Economic Review*, *Journal of Public Economics*, *Rand Journal of Economics* and various other general interest, transportation, industrial organization and regional science journals. In 2001-2002 he was awarded a McCalla Research Professorship from the University of Alberta. His email address is: rlindsey@ualberta.ca.

Econ Journal Watch,
Volume 3, Number 2,
May 2006, pp 380-392.

CHARACTER ISSUES

A List of the 364 Economists Who Objected to Thatcher's Macro Policy

PHILIP BOOTH *

IN THE PREVIOUS ISSUE OF THIS JOURNAL, GEOFFREY WOOD wrote of the important 1981 event when 364 economists signed a petition against Thatcher's macroeconomic policy. Here I offer an appendix listing the signatories. The list includes not only the most eminent economists at the time, but also many people who play a major role in public life today—including the Governor of the Bank of England and one other member of the Bank of England's Monetary Policy Committee (MPC). Indeed, perhaps surprisingly, the signatories include a number of economists who are now Institute of Economic Affairs authors or members of the IEA Academic Advisory Council.

This list is reproduced in the IEA publication, "Were 364 Economists All Wrong?" which can be downloaded from www.iea.org.uk. In the collection, two of the signatories, MPC member Stephen Nickell and Maurice Peston (now a member of the House of Lords), justify the stance they took. A number of other economists explain why they believe that the 364 were wrong and present evidence to justify their case.

* Institute of Economic Affairs and Cass Business School, City University. London, UK SW1P 3LB.

LIST OF SIGNATORIES

The names have been checked against the published list of signatories; however, no checking has been undertaken to ensure that there were no errors in that original list.

Aberdeen University
J. A. Cairns
Professor M. Gaskin
A. H. Harris
A. G. Kemp
P. D. Murphy
D. A. Newlands
Professor D. W. Pearce
R. Shaw

Bath University
Professor D. Collard
R. P. Kamat
P. Mosley

Birmingham University
R. E. Backhouse
R. Clarke
D. G. Dickinson
M. G. Kanbur
Professor M. A. King

Bradford University
Sir Fred Atkinson (Professor)
J. E. Dunworth
M. Wilkinson
P. R. D. Wilson

Bristol University
J. Beath
R. Berry
A. A. Brewer
J. Broome
M. J. Browning
Professor W. H. Buiter
M. Clarke
Professor A. S. Deaton
M. Irish
I. Jewitt
R. Lacomber
H. S. B. Rees
D. C. Webb
D. Winter
L. A. Winters

Cambridge University
T. Barker
I. Begg
V. Borooah
S. Brodersen
Professor G. Cameron
Professor D. Champernowne
K. Coutts
J. C. Craig
F. Cripps
D. A. Dawson
P. M. Deane
J. L. Eatwell
R. Evans
A. Giddens
Professor W. A. H. Godley
A. Goudie
Professor F. H. Hahn
J. Humphries
G. K. Ingham
R. C. Jobling
Lord Kahn (Professor)
Lord Kaldor (Professor)

S. M. Kanbur
M. Kumar
M. Landesmann
A. Lawson
Professor J. E. Meade
G. Meeks
D. E. Moggridge
B. Moore
H. Myoken
Professor R. R. Neild
P. Nolan
M. H. Pesaran
W. Peterson
R. van der Ploeg
Professor W. B. Reddaway
J. Rhodes
Sir Austin Robinson (Professor)
Professor J. Robinson
J. Rubery
P. Ryan
S. Smith
K. Starzec
A. Sutherland
R. Tarling
S. Treitereid
J. Trevithick
Professor H. A. Turner
T. W. Ward
M. R. Weale
J. R. Wells
C. Whitehead
F. Wilkinson

Cardiff, University College
C. Baber
D. Barry
J. S. Bennett
R. Blackmore
T. Boyns
B. Curry

Professor K. D. George
G. Harbour
G. C. Hockley
Professor Sir Bryan Hopkin
C. J. McKenna
R. McNabb
L. Mainwaring
S. Owen
M. Phelps
D. G. Rhys
J. Shorey
D. R. Thomas

City University
J. Ansar
N. Bosanquet
P. Cook
D. Gray
Professor C. D. Harbury
P. Holl
K. R. Kirton
T. E. Tutton

Dundee University
P. G. Chapman
A. A. Lonie
C. M. Lyntag
M. J. Tooze

Durham University
R. A. H. Middleton
P. A. Winston

Edinburgh University
M. Fransman
D. H. R. George
L. T. Oxley
G. C. Reid
C. J. Roberts
S. T. Sayer

Professor T. Vandome

Glasgow University
G. C. Abbott
P. B. Beaumont
N. G. Clark
M. W. Danson
J. Foster
Professor L. C. Hunter
C. Kay
J. L. Latham
W. F. Lever
D. Maclennan
R. G. Milne
E. R. Rado
L. Sirc
Professor A. S. Skinner
Professor A. Slaven

Hull University
Professor J. S. G. Wilson

Keele University
S. A. Hussain

Kent University
A. Carruth
M. A. H. Katouzian
M. T. Sinclair
W. Smith
Professor A. P. Thirlwall

Lancaster University
H. W. Armstrong
V. N. Balasubramanyam
J. Channon
R. W. Daniels
J. Fender
J. E. King
Professor A. I. Macbean

C. Macgregor-Reid
D. T. Nguyen
N. Oulton
D. J. Payne
P. Regan
M. B. Rose
R. Rothschild
P. N. Snowden
J. Taylor
Professor H. Townsend
P. M. Westall

Leeds University
Professor A. J. Brown
J. A. Chartres
H. J. Radice
J. S. Rothwell
Professor M. J. Surrey
Leicester University
D. X. Bhattacharyya
Professor P. M. Jackson

London University
Birkbeck College
F. Atkins
M. Ball
R. Bennett
I. Brunskill
H. Davies
B. Fine
G. Kennally
Professor J. Muellbauer
Professor R. Portes
R. P. Smith
D. Snower
A. Spanos

Imperial College
S. C. Hall
H. Motamen

D. Shepherd
Professor Z. A. Silberston
R. N. Strange
J. N. Turk

London School of Economics
Professor A. B. Atkinson
Professor P. Dasgupta
J. Davidson
C. R. S. Dougherty
J. Le Grande
K. Klappholz
A. Marin
Professor M. Morishima
Professor S. J. Nickell
J. Sutton
Professor P. J. D. Wiles

Queen Mary College
R. J. Allard
J. S. Cubbin
Professor B. A. Curry
H. S. E. Gravelle
Professor M. H. Peston

University College
Lord Balogh
W. Corlett
C. Catephores
C. Heady
Professor P. D. Henderson
A. Markandya
M. Pemberton
K. Schott
Professor J. Spraos
M. Stewart
D. Ulph
R. Vaughan
D. Verry

Manchester University
Professor M. J. Artis
J. M. Currie
P. Devine
M. C. Kennedy
C. H. Kirkpatrick
P. F. Leeson
Professor J. S. Metcalf
T. Peach
W. Peters
D. L. Purdy
J. B. Slater
Professor I. Steedman
P. C. Stubbs

National Institute for Economic and Social Research
P. S. O'Brien

Newcastle University
A. R. Shah

Nottingham University
J. M. Bates
K. A. Ingersent
A. N. Jennings
Professor J. Mitchell
Professor J. R. Parkinson
Professor A. J. Rayner
G. V. Reed
Professor J. H. B. Tew
D. K. Whynes
R. J. Young

Oxford University
B. Banerjee
W. Beckerman
C. J. E. Bliss
Professor Sir Alec Cairncross
A. Graham

V. Joshi
M. C. Kaser
K. Mayhew
A. Oswald
Lord Roberthall
D. Robinson
B. C. Rosewell
Professor A. K. Sen

Queen's University (Belfast)
Professor R. D. G. Black

Reading University
Professor P. E. Hart

Sheffield University
R. Clarke
Professor G. Clayton
D. J. Goacher
D. Kitchin
R. Lawson
J. Peirson
S. G. Tebbutt

Southampton University
C. J. Hawkins
Professor K. Hilton
G. W. McKenzie
Professor G. E. Mizon
R. J. O'Brien
Professor D. C. Rowan

Stirling University
P. J. W. N. Bird
Professor C. V. Brown
M. S. Common
G. Evans
D. Ghosh
R. T. Hamilton
P. G. Hare

C. Normand
R. Shone

Strathclyde University
R. G. Brooks
Professor A. I. Clunies-Ross
K. Hancock
J. Sconller
P. Wanless

Sussex University
Professor T. Barna
P. Lesley Cook
P. Holmes
D. Hunt
Professor D. Winch

Swansea, University College
N. Baigent
D. Donneky
J. T. Harris
L. C. Hunt
I. Jeffries
H. C. Petith
D. E. L. Thomas

University of East Anglia
D. Bailey
Professor K. N. Bhaskar
S. W. Davies
M. Dietrich
A. E. B. Heading
Professor H. F. Lydall
Professor A. Parikh
Professor A. G. Schweinberger
J. T. Thoburn

University of Wales, Bangor
Chakravarty
J. Fletcher

E. P. Gardener
R. H. Gray
A. J. Hopkins
R. R. Mackay
Morrell
Professor J. R. S. Revell

University College of Wales, Institute of Science and Technology
N. F. B. Allington
P. R. Bridger
Professor Glyn Davies
J. W. Evans
G. M. Holmes
T. W. Taylor
D. R. Turner

Warwick University
J. Brack
P. Burridge
Professor K. Cowling
Professor A. G. Ford
M. Harrison
C. G. Hayden
N. J. Ireland
K. G. Knight
P. J. Law
D. Leech
Professor M. Miller
S. K. Nath
G. Renshaw
J. I. Round
B. Sadler
M. Salmon
A. Snell
Professor N. H. Stern
M. Stewart
P. Stoneman
P. A. Weller

York University
D. Austen-Smith
R. R. Barnett
P. Burrows
Professor C. H. Feinstein
K. Hartley
J. P. Hutton
D. T. Jenkins
P. J. Lambert
J. M. Malcolmson
A. K. Maynard
D. S. Poskitt
J. W. Posnett
M. Sawyer
P. M. Solar
J. Suckling
G. B. Stafford
R. B. Weir
Professor A. Williams

ABOUT THE AUTHOR

Philip Booth is Editorial and Programme Director at the Institute of Economic Affairs and Professor of Insurance and Risk Management at Cass Business School, City University. He has edited the IEA monograph *Were 364 Economists All Wrong?* He is also author of a number of books and papers in the fields of pensions, social insurance, investment and finance. From 1998-2002, Philip Booth was a special adviser to the Bank of England. His email address is: pbooth@iea.org.uk.

GO TO WOOD (2006) ARTICLE ON THE 364

Econ Journal Watch,
Volume 3, Number 2,
May 2006, pp 393-397.

INTELLECTUAL TYRANNY OF THE STATUS QUO FOLLOW UP

Reply to Hortlund's "Defense of the Real Bills Doctrine"

RICHARD H. TIMBERLAKE*

IN A CRITIQUE OF MY PAPER ON "GOLD STANDARDS AND THE Real Bills Doctrine in U.S. Monetary Policy" that appeared in *Econ Journal Watch* (August 2005), Per Hortlund has raised several interesting issues about the Real Bills Doctrine (RBD). As Hortlund observes, my article had two major themes, first, the innocence of the gold standard for the monetary infelicities that caused the Great Contraction of 1929-1933, and, second, the culpability of the RBD for the debacle. Hortlund accepts my defense of the gold standard. However, he finds some arguments to support a case for the RBD, and he raises an important issue concerning the substance of the RBD and its implementation as policy during 1929-1933, part of which I thoroughly accept.

Hortlund states that the RBD "may be a useful rule of thumb that can improve the workings of a central banking system on the gold standard" (74). No one would deny this assertion. The more fundamental question is, however, would the RBD by itself—without a gold standard to protect it—have the necessary operational constraints to prevent dynamic instability in the monetary system? I do not find Hortlund's affirmative arguments on this particular question convincing. His next sentence noting that the RBD prohibits the monetization of government debt is correct and

* Department of Economic (retired), University of Georgia. Bogart, GA, 30622.

attractive even though the idea may have come about because of the real billsers' hostility to the quantity theory of money.

Hortlund argues, I think mistakenly, that the RBD can be made as rigorous as a gold standard:

> Note that the real bills criterion in this sense is a purely formal, legal one—similar for example to the criterion of monetization of gold on fixed terms. Nothing says that a loan bill of a certain nominal amount could not represent "real value" or "real wealth". Timberlake's (2005, 206) discussion about the distinction between monetization on gold and on real bills I therefore believe to be somewhat off the mark. (Hortlund 2005, 76)

But how does his "loan bill of a certain nominal value" come to represent real wealth? The valuation of real bills always requires human discretion by bankers or central bankers. At times, bankers' judgments may be stabilizing, and at times destabilizing.[1] By way of contrast, the monetization of gold under a gold standard can never be dynamically unstable. Gold standard laws regulate the monetization of gold in a way that laws describing and governing real bills can never achieve.

I realized after writing "Gold Standards and the Real Bills Doctrine in U.S. Monetary Policy" that I had neglected to emphasize sufficiently the negative features of the RBD; Hortlund's "Defense" provides this proper amendment.

Virtually all of the arguments on the workability of the RBD focus positively on monetization of the financial assets that qualify as real bills. However, in 1929 the Fed Board instituted its hysterical crusade against "speculation," which was very much a negative policy. Other negative types of lending that the RBD considered unsuitable were long-term investments and real estate, and, as Hortlund emphasizes, government bonds.

I wrote in a short sequel to "Gold Standards . . .,"

> The Fed Board's anti-speculative compulsion crowded out legitimate lending to needy banks that actually had 'real bills' to discount, thereby causing the very condition that the Federal Reserve System was supposed to prevent. (Timberlake, unpublished)

[1] For an analysis of the conditions for both, see Girton (1974).

Hortlund takes the matter further.

> The true real-bills central bank discounts *if and only if* real bills (eligible paper) is presented to it for discount. That is, the real-bills central bank acts according to these two rules:
>
> Rule #1: Do not discount non-real bills (ineligible paper) presented for discount.
>
> Rule #2: *Do* discount all real bills (eligible paper) presented for discount.
>
> A central bank that does not follow *both* of these rules cannot be said to follow the RBD. For the RBD adherent, the real bills criterion provides a rule that renders monetary policy more or less automatic. The volume of real bills corresponds by itself in a desired manner with the 'needs of trade', wherefore no discretionary evaluation or prudence on the part of the central bank is needed. The central banker's only task is to check whether the instruments presented for discount is on the list of eligible paper or not. Did the Fed follow the rules? (Hortlund 2006, 82).

The obvious answer, of course, is "No." However, the Fed's failure to discount real bills at this time, due to the anti-speculative bogeyman it had fabricated, does not confirm that it had abandoned real bills precepts. Indeed, "speculation" is a major sin in real bills circles. It is just as much a part of the RBD as are the positive principles for lending. Consequently, one cannot separate the anti-speculative proscriptions that the Fed pursued during 1929-1933 from the RBD's proper prescriptions. Both together make up the total package. That being the case, the disaster of 1929-1932 still ranks as the result of the RBD* (with an asterisk signifying that the negative side of the Doctrine was dominant).

One cannot argue that the contraction never would have occurred if Fed policymakers had just discounted real bills. It is a counterfactual impossibility. They did not do so because their real bills precepts required them to eliminate speculative 'credit' first. Fed policy expanded Hortlund's

Rule #1 of not discounting non-real bills to an active negative policy of killing speculative 'credit' no matter what the cost.

Incidentally, if I could set the rules for a central bank, in addition to Hortlund's two rules, I would add a third rule.

> Rule #3: Do not pay any heed to any particular market, such as the stock market, the real estate market, the agricultural market, the foreign exchange market, or any special asset or capital problems, or any goal or target at all, except stability in the value of the money unit. Use your absolute control over the quantity of money to achieve this goal, and with zero tolerance.

The crux of this Rule is that the mechanics of contemporary central banking give that institution control over one, and only one, Big Variable—the quantity of money. Given this one control, the central bank can maintain stability of the money unit. However, in doing so, since it cannot aim at multiple targets at the same time, it must eschew all other considerations—no matter how much its governors discuss them or wring their hands about them. In short, both the necessary and sufficient condition for managing a central bank in today's world is the quantity theory of money.

Benjamin Strong understood this principle and acted accordingly. Indeed, his appreciation of the Fed's power, as a part of the government, to control the quantity of money and the price level is what prompted him to reject a stable price level mandate for Fed policy. He wanted permanent control of money ruled by a gold standard, with the central bank limited to providing some currency 'elasticity' in a crisis. The best we can hope for from the present-day Fed is something akin to Strong's policies.

REFERENCES

Girton, L. 1974. SDR Creation and the Real-Bills Doctrine. *Southern Economic Journal* 41(1): 57-61.

Hortlund, Per. 2006. In Defense of the Real Bills Doctrine. *Econ Journal Watch* 3(1): 73-87. Link.

Timberlake, Richard L. 2005. Gold Standards and the Real Bills Doctrine in U.S. Monetary Policy. *Econ Journal Watch* 2(2): 196-233. Link.

ABOUT THE AUTHOR

Richard H. Timberlake, Jr. Born June 24, 1922, in Steubenville, Ohio. Served in World War II. Attended Kenyon College (B.A., 1946), Columbia University (M.A., 1950) and the University of Chicago (PhD, 1959). Taught economics at Muhlenberg College, Norwich University, Renssalear Polytechnic, Florida State University, and at the University of Georgia for 26 years. Retired in 1990. Married to the former Hildegard Weber of Essen, Germany, and have five children. Research has been primarily on the history of money and of monetary policy. Principal publications: *Money and Banking* (with Edward Selby), Wadsworth, 1972; *Origins of Central Banking in the United States,* Harvard University Press, 1978*; Monetary Policy in the United States*, Univ. of Chicago Press, 1993; Co-editor, *Money and the Nation State*, Independent Institute, 1998. Current research: Analysis of the monetary decisions of the Supreme Court. Address: 220 Deerfield Road; Bogart, GA 30622. E-mail: rightim@earthlink.net.

GO TO ORIGINAL TIMBERLAKE (2005) ARTICLE

GO TO HORTLUND (2006) COMMENT ON TIMBERLAKE