

MICROECONOMIC TESTS OF RICARDIAN EQUIVALENCE

James P. Cunningham¹

Chapman University

and

Arnold C. Harberger¹

University of California, Los Angeles

August 2005

ABSTRACT

This paper tests the Ricardian Equivalence Concept by constructing “Ricardian” time series for the incomes of U.S. States. “Ricardian” income is taken to be standard disposable income minus the net increment to state debt plus the State’s apportioned share of Federal Debt. Regressions are then run using “Ricardian” and, alternatively, “standard” concepts of disposable income to explain residential electricity consumption in each state. The simplest tests used just real income and relative price as regressors. Later tests added weather and natural-gas-price variables. The “standard” concepts prove preferable, at high levels of significance in all the tests performed.

* * * *

The purpose of this paper is threefold: a) to make a further contribution to the already substantial literature on the empirical testing of hypotheses concerning Ricardian Equivalence, b) to demonstrate the value and illustrate some of the subtleties of working with subnational (in this case U.S. State) data, and c) to advertise the existence of a new data set on the public finances of the 50 states, which may be of considerable use to others in our profession.

Some Thoughts on Ricardian Equivalence

Ricardian Equivalence follows quite "standard" economic principles in viewing individual and household behavior as resulting from a process of maximizing the entity's utility subject to a "lifetime budget constraint." The basic principle underlying the proposition is that the pattern of consumption over time should not respond to disturbances or modifications of initial conditions that leave unchanged the present value of wealth (i.e., the present value of the lifetime budget constraint). In particular, consumers should not only be indifferent as between a) paying an additional \$1000 in taxes now and b) paying over future periods a pattern of incremental taxes whose present value is \$1000; but, in addition to their being indifferent, their solution to the problem of distributing their planned consumption ($C_1, C_2, C_3...$) over time should also remain unchanged. From this it follows quite directly that if consumers respond to an immediate extra tax of \$1000 by contracting present and planned consumption in a specified way, they should also respond in the same way to the government's issuing \$1000 extra debt instruments to finance the same outlays. In the simplest scenario, consumers willingly buy the bonds the government just issued, and hold them until the taxes to service or pay them are levied. The implicit assumption here is that future taxes with an expected present value of \$1000 will have to be levied to pay (or otherwise service) the increment. This is easy to see if the debt is fully paid in one given future year j , where the amount needed to pay off the debt is simply $\$1000(1+r)^j$. This obviously has a discounted value of \$1000; but so also does a perpetual payment of interest at $\$1000r$ per year. Clearly, at this level of abstraction, it should not matter when, or in what time pattern, or even whether the \$1000 increment of debt is paid, so long as the government does not default on either the interest or principal along the way.

From the moment of its revival in the mid-1970s [see Barro (1974)], the idea of Ricardian Equivalence has provoked controversy. No one, to our knowledge, has disputed the internal logic or consistency of the type of exercise presented above; instead, critics produced a long list of reasons why real-world cases would differ in important ways from the explicit or implicit assumptions of the exercise. [For a review of the literature, see Seater (1993).]

Perhaps the simplest case for discussing Ricardian Equivalence is that in which the \$1000 increment of bonds is sold to (and held by) foreigners. In that case the comparison for the domestic economy is simply one of "tax now" or "tax later". Here the criticism boils down to two simple possibilities -- a) domestic consumers have a full and clear appreciation of the burden of extra future taxes needed to service and/or pay the debt, but discount those expected future taxes at a different, presumably significantly higher rate than r , and b) regardless of whether domestic consumers use r as their discount rate, they do not fully incorporate the expected future tax payments in their current calculations of their lifetime budget constraint.

These extreme possibilities are not mutually exclusive; in reality the combination of incomplete perception and a higher discount rate is much more plausible than either alone. Indeed, if people react to future uncertainties by applying a higher discount rate, a fuzziness of perception of future tax payments might have its reflection not in a reduction of type b) in their perceived value, but instead in an increase of type a) in the rate used to discount them.

The above case stays quite close to the theoretical framework of Ricardian Equivalence, simply inserting different behavior assumptions into the framework. An alternative case would maintain a firm tie to consumer rationality as a principle, but would modify the way in which economic agents are postulated to view the future. Under this alternative, people will likely follow standard rationality assumptions in making certain major intertemporal decisions (life and

disability insurance, retirement plans, funding children's education, etc.), but do not perceive clearly what they cannot "see" clearly. [For survey data of economics students' (lack of) knowledge of the debt, see Gruen (1991).] Instead of reacting consciously to a great multitude of vaguely possible future contingencies, they instead think of their own physical and financial assets as a cushion against all such poorly-perceived contingencies. According to this view, people really do take expected future tax liabilities (and other contingencies) into account in their decisionmaking, but they make their adjustments only or mainly after a certain threshold of likelihood is crossed. For example, property tax bills come twice a year in many localities, with the first bill (say in February) being a good but not perfect predictor of the second bill (say in August). It is quite certain that in these circumstances most consumers will take the February bill as a signal of a roughly similar August bill to come, and will incorporate both liabilities in their calculations as of February.

It is not the same when the precipitating event is an increment of government debt. Here the link to future tax liabilities of any given individual or household is very weak indeed. It is unlikely, therefore, that households will respond to the increment of debt itself. Instead they will likely wait for more direct signals with respect to their future tax payments. The increase in debt is a very remote signal; the discussion in the press (probably much later, and probably unrelated to any given increment of debt) and in political circles of the need for a tax increase is still remote, but less so; the development of competing plans for tax changes comes closer to home; specific bills introduced in Congress come even closer, their legislative passage brings them still closer, and signature by the President brings them closer yet. But probably most taxpayers will make only modest, partial adjustments at the different stages of this process. The main adjustment will come at the point where the household has a pretty good idea of how the new

legislation will affect its own members. This may be as early as the time when the competing tax plans are discussed (especially if the household would fare much the same under all the likely alternatives), or as late as the time when the household actually fills out its tax returns under the new legislation (particularly if the new legislation is typically complex, maybe even with provisions that operate in offsetting directions for some households).

This position, to which we ourselves lean, does not imply that households are irrationally myopic, only "living for the present." It fully permits the household to spread its response to a new tax over all future years, even when it thinks the new tax will only last for a specified period. It postulates full or close-to-full rationality on the part of decisionmakers as soon as information is sufficiently clear as to enter their "circle of light," where decisions are made. Information that is as yet too vague and inchoate to enter this circle remains in the surrounding "penumbra of darkness," against whose multiple but vague contingent threats households try to protect themselves by broad, general strategies such as the progressive accumulation of wealth.²

Further complications are added to the discussion of Ricardian Equivalence when it is assumed that the incremental debt is held by domestic households rather than by foreigners. At this point the issue arises of whether the holders of the debt are the same households as the actual or prospective taxpayers. Representative consumer models assume them to be the same, but empirical studies of who holds the debt support a quite different view. It is totally appropriate, in our view, for households that hold government bonds to consider those bonds to be fully and completely an asset, and the interest payments to be fully and completely income flows to the household. Thus, there should in our view be no debate as to whether government bonds are gross wealth to the households that own them. Whether they are "net wealth" or not depends

upon the behavior of all (not just bond-holding) households vis-a-vis the future tax liabilities that those bonds represent.

Under these circumstances Ricardian Equivalence may fail to hold if there are important differences between bondholders on the one hand and the generality of households on the other, either with respect to the discount rates they use or with respect to the clarity with which they perceive the specific future flows in question. On discount rates, it is to be presumed that households with internal discount rates higher than \underline{r} will almost by definition not be bondholders, while those who do hold bonds should have internal discount rates that are less than or equal to \underline{r} . The perception of future flows is similarly much less of an issue with respect to bondholders (who can presumably at any time sell their bonds at the market) than it is with respect to potential future taxpayers, whose position is not really any different from what we have already discussed for the case where the bonds were held abroad.

We see a presumption, therefore, in favor of the proposition that households do have different perceptions as between taxes and increments of government debt. This is a presumption against strict Ricardian Equivalence.

There is a further element that could well operate in an important way to cloud a comparison of the apparent effects of taxes versus increments to government debt. This has to do with how the money is spent. Up to now, our discussion has centered on taxes versus new indebtedness as alternative ways of financing a given set of expenditures. At the opposite extreme we could work on the assumption that new indebtedness was the natural source of financing for capital expenditures, while taxes played this role for current expenditures and for debt service that had to be paid out of general revenues. This assumption is certainly extreme vis-a-vis the realities of our time (or of our world), but it has a sensible economic foundation, in

that it can be thought of as consciously attempting to shift to future generations of consumers the capital costs associated with the public services they will enjoy.

In this scenario, the Ricardian proposition that an increment of debt acts "just like taxes" takes a quite different form. In the standard presentation of equivalence taxes are unequivocally "bad", reducing consumers' utility and welfare. When an increment to debt is enacted in lieu of a certain amount of taxes, people are posited to treat that increment as being equally "bad", because it represents the present value of taxes yet to come. In contrast, the new scenario would have people being neutral vis-a-vis an increment to debt, with prospective future costs, whether in the form of tolls, utility rates, or general taxes, being roughly matched by prospective future benefits. Strict Ricardian behavior would entail consumers' also being neutral to tax changes "in lieu of" debt increments--i.e., tax changes that financed capital acquisitions that promised future benefits, but not to all tax changes (assuming taxpayers still treat the bulk of taxes simply as costs).

This line of reasoning provides a basis for expecting consumers to react more "neutrally" to increments of debt than they do to taxes. Hence it provides a "Ricardian" rationale for the observation that consumption variations are better explained by a standard disposable income variable than by one that is adjusted for changes in public debt. So supporters of Ricardian Equivalence have at least one basis on which they might conceivably rationalize such an observation. We feel, however, that the separate links in this chain of reasoning are too weak to endow it with serious credibility in explaining the overall reactions of consumers to increments of debt. The standard Ricardian argument has consumers viewing increments of debt as measuring taxes yet to come. This modified argument has increments to debt being perceived by consumers as measuring taxes-cum-benefits yet to come. Both cases posit an acuity of

perception with regard to distant future events and contingencies that we find highly implausible.

The Nature of Our Experiment

We plan, using state data, to test whether people make their spending decisions more in response to changes in disposable income as conventionally defined, than in response to a new variable--disposable income minus the increment of public sector debt that is allocable to them. This new variable treats increments to debt as "taxes in disguise", which is precisely how they are viewed under Ricardian Equivalence.

Our tests will be carried out by allowing these two definitions of income to compete with each other, to determine which of them better explains the observed facts. An alternative would have been to introduce the increment of debt as a separate variable, alongside disposable income, allowing it to have its own, separate coefficient. We rejected this mainly because Ricardian Equivalence, in its standard form, asserts that taxes and increments to debt operate in the same way, and hence should be treated in the same way (in our case, merged into a Ricardian disposable income variable). But also, we were aware that the "increment to debt variable" would be small, relative to disposable income, and also quite volatile. We would not expect it to have a statistically significant coefficient in most cases, which would mean that whoever designed the experiment could virtually determine the result.

That is to say, a Ricardian Equivalence person could define income by subtracting increments to debt as well as taxes, and then test increments to debt as a separate variable, while a traditionalist would define income (for regression purposes) by subtracting just taxes, once again treating increments to debt as a separate variable. The presumption is that in both cases the

separate "increments to debt" variables would be statistically insignificant: hence, whichever null hypothesis one chose to start out with would end up being sustained.

To surmount this problem, we chose to allow two income definitions to compete head-to-head; and to place our primary reliance on a simple binomial test based on which of the two income variables yielded the higher R^2 (smaller sum of squared residuals) in regressions that were identical except for the income variable. Using separate regressions for the 49 states of our study (Maryland data were incompatible with the rest, because they were grouped with District of Columbia data), this simple binomial test had the potential of yielding significance levels as high as $(1/2)^{49}$, even if not a single one of the 49 differences in R^2 was itself significantly different from zero!

We do not actually have cases in which we observe 49 "heads" and zero "tails". But we do report on some exercises in which the non-Ricardian hypothesis produces the higher R^2 in as many as 44 cases being compared, a result that could occur by chance alone less than one time in a trillion!!!

Explaining Household Electric Energy Consumption

Once the decision was made to test Ricardian Equivalence by setting up a contest between two alternative income variables, it was but a small step to explore possible behavior relations in which income played an important role. Standard demand equations came immediately to mind, and opened a whole new continent for testing Ricardian Equivalence. Within the vast array of possibilities, we quickly focused on household electricity consumption. [For a discussion of how the selection of the consumption variable affects Ricardian tests, see Graham (1992).] Electricity at the level of household consumption is about as homogeneous as a good can get. It is also consumed by nearly all households, with excellent data on both quantity

and value being available at the state level. Moreover, state boundary problems, which create horrible snares in dealing with demand for alcoholic beverages or tobacco products, are virtually non-existent except for Maryland (whose data are mingled with those of Washington, D.C. for much of the period). Our observations cover the period 1963-93 for all U.S. states except Maryland.

We concluded very early that we should avoid all temptation to aim at one grand test, on the basis of pooled cross-section and time-series data. We found no reason at all to think that a single overall demand function would prevail in all the different states. Demands for heating, air conditioning and lighting vary greatly by region, and each of these uses presents quite a different situation with respect both to price and income elasticities. We note, for example, that close substitutes (coal, oil, gas, even kerosene in some cases) are almost universally available in the case of heating uses. For air conditioning, only gas comes close to performing this function; for lighting there simply is no serious substitute. So we expect significant variations in price elasticity, as among the states. As far as income elasticity is concerned, one would run into a mare's nest of problems trying to work with panel data. Differences in heating and air conditioning demand across the states have a huge "permanent" component, reflected among other things in the size and architecture of dwelling units. A regression based on panel data would try to explain these differences side by side with the time series variations within each state. We reasoned that those variations would have little to do with the income concepts being tested. On the other hand, if we just think in time series terms for the individual states, the presumption is that the underlying income elasticities will, like the price elasticities, be very different owing to the different nature of the underlying demand. Air conditioning demand has

surely, over our period of observation, been much more responsive to income growth than has heating demand and, presumably, lighting demand.

Having decided on individual regressions for the 49 states, and on a head-to-head confrontation of two competing definitions of income, we then faced the question of how to set the rules of the race. Many considerations entered here. First, we settled on a logarithmic form for the quantity, income and price variables, mainly because it was so much easier and more natural for us to think and work (and communicate our thinking and results to others) in terms of elasticities than in terms of slopes. Elasticities also provided an easy way of comparing results across the 49 states.

Second, we wanted our work to reflect the main alternative representations of income. To this end we settled on two alternative definitions: current year real disposable income on the one hand, and a 4-year weighted moving average of real disposable income on the other. The latter variable was intended to reflect the considerations that lie behind the permanent income hypothesis and its siblings.

We may have departed from tradition, however, in insisting on imposing the weights (we chose 0.4 for current real disposable income, and 0.3, 0.2, and 0.1 for the three successive lagged values of the same variable), rather than allowing them to be chosen by the data (either through introducing separate lagged income variables or through letting the data choose a parameter describing the rate of exponential decay in the weighting function). We are suspicious of the first procedure because of the erratic patterns of lagged coefficients that it quite often produces. More importantly, we feel that both procedures depart from the underlying rationale for concepts like permanent income. That rationale uses permanent income as something like a proxy for the "perceived" lifetime budget constraint. This means that there should in principle be just one

correct measure of permanent income. But getting a different weighted coverage for each commodity whose demand equation one estimates takes off in a very different and contradictory direction.

Similarly, though we are very ready to recognize that price and income elasticities of demand for electric energy should and will exhibit significant variation across states and regions, we find little ground for thinking that such variation would characterize how American consumers from the different states behave in assessing what is their "permanent income". The choice of 0.4, 0.3, 0.2 and 0.1 as the weights applied successive lags of real disposable income is arbitrary but sensible (using annual data). The basic idea of the moving average is to allow past incomes to have some weight. Putting two thirds or three quarters of the weight on current income would almost belie this basic idea. On the other hand, a weighting pattern stretching back, say, a decade, and built on the same sum-of-the-years-digits principle, would give less than 20% weight to current income, and more than a quarter to (the sum of) incomes going back 5 to 10 years. This, too, seems grossly implausible. The chosen weighting pattern lies in the plausible range; it is easy for everybody to understand; and, finally, it gives sensible and plausible results when applied.³ Harberger and his students have used it in many applications with good success. They have also found that the results are quite insensitive to plausible variations in this lag structure.

The third key decision concerned imposing limits on the acceptable range for key parameters. We know certain key facts about electricity demand--that it has in the past been quite responsive to variations in real income, and somewhat less responsive to changes in its own relative price. These considerations led us to impose an income elasticity of 0.25 in cases where the free estimate was less than that number, and a price elasticity of -1.5 in cases where the free

estimate exhibited an even higher elasticity. We also imposed an upper limit of 1.5 on the income elasticity.

The only other cases of imposed parameters derived from the Slutsky conditions: an own-price elasticity of zero was imposed in cases where the free estimate was less than zero, and a cross-elasticity of zero with respect to the relative price of gas was imposed in cases where its free estimate was negative (i.e., we ruled out complementarity between the residential uses of electricity and gas). All price elasticities are compensated ones, as income is invariably measured in real terms. As a check on the reasonableness of these presumptions, we note that the great bulk of the free estimates of income elasticities were between 0.4 and 1.25, while those of price elasticity were highly concentrated between zero and -0.8.

Equations, Results and Statistical Tests

We here present results of four pairwise comparisons of estimated demand relations. Within each comparison the only difference is the shift from a traditional disposable income concept to a Ricardian disposable income concept. We define:

Q = residential electricity consumption per capita in a given state (in Mwh).

Pe = total billings for residential electricity in the state divided by QN (N = population of the state) and deflated by the GDP deflator.

P_g = average price within the state of 1000 cu. ft. of natural gas at standard temperature and pressure, deflated by the GDP deflator.

y_d = personal disposable income per capita in the state deflated by the GDP deflator; all for year t

y_r = Ricardian disposable per capita in the state ($= y_d - \Delta D$), where ΔD for a state consists of the net flow of funds to that state on debt account ($=$ new issues minus retirement),

during year t, deflated by the GDP deflator and expressed in per capita terms.

$$y_{4dt} = .4 y_{dt} + .3 y_{d,t-1} + .2 y_{d,t-2} + .1 y_{d,t-3}$$

$$y_{4rt} = .4 y_{rt} + .3 y_{r,t-1} + .2 y_{r,t-2} + .1 y_{r,t-3}$$

C_t = cooling degree days in the state in year t -- i.e., the number of degrees by which mean daily temperature exceeded 65° F, summed over all days of year t.

H_t = heating degree days in the state in year t -- i.e., the number of degrees by which 65° F exceeded mean daily temperature, summed over all days of year t.

The fitted equations were:

$$(1a) \quad \ln Q_t = \text{const.} + \varepsilon_d \ln y_{dt} + \eta \ln P_{et}$$

$$(1b) \quad \ln Q_t = \text{const.} + \varepsilon_r \ln y_{rt} + \eta \ln P_{et}$$

$$(2a) \quad \ln Q_t = \text{const.} + \varepsilon_d \ln y_{4dt} + \eta \ln P_{et}$$

$$(2b) \quad \ln Q_t = \text{const.} + \varepsilon_r \ln y_{4rt} + \eta \ln P_{et}$$

$$(3a) \quad \ln Q_t = \text{const.} + \varepsilon_d \ln y_{dt} + \eta \ln P_{et} + \sigma \log P_{gt} + \lambda C_t + \mu H_t$$

$$(3b) \quad \ln Q_t = \text{const.} + \varepsilon_r \ln y_{rt} + \eta \ln P_{et} + \sigma \log P_{gt} + \lambda C_t + \mu H_t$$

$$(4a) \quad \ln Q_t = \text{const.} + \varepsilon_{4d} \ln y_{4dt} + \eta \ln P_{et} + \sigma \log P_{gt} + \lambda C_t + \mu H_t$$

$$(4b) \quad \ln Q_t = \text{const.} + \varepsilon_{4r} \ln y_{4rt} + \eta \ln P_{et} + \sigma \log P_{gt} + \lambda C_t + \mu H_t$$

All of our empirical tests focus on comparing an equation of type a) with its "partner" of type b). The first set of tests simply counts how often a given variant [the traditional (a) or the Ricardian (b)], yielded the higher adjusted R^2 out of N pairwise comparisons. The binomial distribution is then used to determine levels of significance. This set of tests has the virtue of great simplicity and easy communicability to the results. The only limitation that we perceive is

that it gives the same points to the winner of a race, independent of whether the victory was "by a nose" or "by a mile."

We pursue two alternative routes to get around this problem. The first of these is very easy to visualize and involves simply considering "close" victories as if they were ties. The results of following this route are presented in Table 1. The table consists of four panels, each of which represents a specific regression format. Panels 1 and 2 are derived from regressions in which per capita residential energy consumption is treated as a function simply of real per capita income and of the relative price of residential electricity. Within this pair, panel 1 is based on regressions 1a and 1b, which use current income as the (traditional or Ricardian) income variable; while the regressions for panel 2 follow equations 2a and 2b, and use an expected income concept for both the alternatives being compared.

Panels 3 and 4 are based on more elaborate demand equations, incorporating the relative price of gas, the number of heating degree days and the number of cooling degree days as additional regressors, following equations 3a, 3b and 4a, 4b respectively. As was the case between panels 1 and 2, panels 3 and 4 differ only in terms of their use of the current income concept (panel 3) or an expected income concept (panel 4) in their definition of the income variable.

It can be readily seen that the non-Ricardian alternative always registers more victories (i.e., more states in which it has the higher adjusted R^2) than its "Ricardian twin". In panel 1 the non-Ricardian definition completely blanks the Ricardian one, except for a single case where the Ricardian alternative wins "by a nose", in just one state. But this Ricardian victory disappears as soon as even the most narrow definition of a tie is introduced. After this point, no further Ricardian victories are registered in that panel.

In panel 2, one starts with 42 non-Ricardian wins, as against 3 Ricardian wins. As different definitions of a tie are introduced, the number of Ricardian victories remains constant at 3, while the number of non-Ricardian victories stays mostly between 39 and 42, dropping to 36 only under the widest definition of a tie.

In panel 3, one starts with 38 non-Ricardian wins, versus 10 for the Ricardian alternative. The number of ties grows faster here, but the ratio of non-Ricardian to Ricardian wins grows as one moves down the table, ending at 18/2 for the most generous definition of a tie.

Panel 4 shows a similar growth in the number of ties as one moves to broader and broader definitions of a tie. But here the tendency is for the number of non-Ricardian wins to always stay around 3 times the number of Ricardian victories. The “worst” case for the non-Ricardian approach is the last line of panel 4, with 12 non-Ricardian versus 4 Ricardian urns. But even this is far from trivial. The chance of getting 4 or less heads in 16 flips of a coin is almost exactly 4%, that of getting 4 or less of either kind is a little short of 8%.

There can thus be no doubt with respect to who is the victor in Table 1. It is the non-Ricardian hypothesis all the way.

Table 2 reflects the results of the more sophisticated comparison of the two alternative income definitions. It can most easily be visualized as the outcome of a sequential process, each “draw” being represented by a pair of regressions explaining residential electricity consumption in a given state. The method can be thought of as estimating at each step the “odds” favoring one or the other of the two “explanations”. Since “odds” are a ratio of probabilities, we can say that if at the end of the first draw the odds were 2 to 1 in favor of hypothesis a, this means $P_{1a}/P_{1b} = 2$, where P_{1a} and P_{1b} are the probability levels emerging from that draw. The probability level derived from the second draw will then be P_{2a}/P_{2b} , but the cumulative odds,

incorporating evidence from both draws, would be the product of the two ratios i.e.,

$(P_{1a}/P_{1b})(P_{2a}/P_{2b})$. Where there are k draws, we have the cumulative odds equal to

$\prod_{j=1}^k (P_{ja}/P_{jb})$. This simply builds on the familiar proposition that the joint probability of a

sequence of independent events is the product of their individual probabilities. The last piece of

the puzzle is the proposition that $(P_{ja}/P_{jb}) = s_b^2 / s_a^2$ -- that is, the odds favoring a given

hypothesis are inversely proportional to its mean squared residual.

Assessing the Plausibility of the Results

There can be no doubt as to the statistical significance of our results, but we consider it important to probe into their economic plausibility as well. Table 3 present data on the distribution of estimated income elasticities of residential electricity demand, and Table 4 does the same for the estimated own-price elasticities.

Some Technical Details

If the course of fitting the various regression equations, we encountered in a small number of instances the problem of a mean of calculated squared residuals that was greater than the variance of the original dependent variable--i.e., cases of "negative R^2 ". This was not really surprising, because in each such case regression coefficients had been imposed as a result of the previously-mentioned restrictions on income and price elasticities. To deal with such cases in our first set of tests, we adopted the convention of disqualifying the contestant [(a) or (b)] in question. If only one contestant was disqualified, the other was declared the winner in the binomial race. If both were disqualified, the pairwise comparison in question was eliminated from the sample.

To deal with the same problem in our second set of tests, we opted for substituting the variance of $\log Q$ for the variance of the residuals in cases where the latter was greater. If, for a given state, both alternatives [(a) and (b)] fell in this category, that automatically imposed a variance ratio of unity for that state. If only one of a) or b) was thus affected, that income concept was automatically the loser for that state. But the other of the two, the winner, was given points only to the extent that it beat the observed variance of $\ln Q$. We did not want to give it extra points that came directly from the restrictions we had imposed, in cases where (however plausible those restrictions might be) they led to a residual variance that exceeded the original variance to be explained.

Our “Best” Tests: Using Seemingly Unrelated Regression Procedures

The next iteration represents the end result of a series of steps, aimed at coping with a number of potential “loose ends” in our estimation process. Our first step was to check for autocorrelation; here Savin and White’s (1977) test for autocorrelation was applied, leading to the rejection of first-order serial correlation. (The actual coefficient of correlation between the residuals u_t and u_{t-1} , for a pooled regression was 0.077 for equation 4a and 0.092 for equation 4b.)

Second, we tested for heteroscedasticity by applying White’s (1980) test. This test was passed at a probability level of 0.50.

Third, we assessed the independence of the residuals across the different states for which our regressions were fitted. Here the Lagrange multiplier test of Breusch and Pagan (1980) indicated that the residuals were not sufficiently independent to be consistent with diagonality of their variance/covariance matrix.

The relevant statistic is $\lambda = T \sum_{i=2}^M \sum_{j=1}^{i-1} r_{ij}^2$. Its level, for our regression 4a, for example,

was 2933.47, compared to a critical value of 163.68 (for $df = 120$). Here the r_{ij} are the individual correlations between the residuals of the regressions for states i and j , and T is the number of observations in the individual regressions (in our case 32, for the years 1963-1994).

This is the route that led us to apply Zellner's technique of seemingly unrelated regression (SUR) models, which makes, in our view, the most discriminating use of the data. In applying SUR, we did not constrain the parameters in any way, allowing each state to have its own constant term, its own price and income elasticities, etc. The technique did, however, allow for common elements to be reflected in the error terms of any and all pairs of included states.

We did run into a snag, however, as we carried forward the SUR exercise -- namely, a variance/covariance matrix of state residual terms that approached singularity if all 49 of our states were included. Our diagnosis was that the determining variables and residuals of some pairs (or combinations) of neighboring states were such that one member of the pair could hardly be distinguished from another (or some linear combination of others). We coped with this problem by progressively eliminating states from the comparison, until the problem of singularity disappeared. Our aim in this winnowing process was to drop small states rather than large ones, and to maintain a reasonable representation for each geographic region.

Thus we have Maine, New Hampshire, and Massachusetts but not Connecticut, Rhode Island and Vermont; New York and Pennsylvania but not New Jersey and Delaware; Virginia, West Virginia and Kentucky but not Tennessee; Texas but not Oklahoma; Washington but not Idaho; Montana but not Wyoming; etc. In all, the left-out states accounted for only 19% of total residential electricity consumption.

Significance and Plausibility of SUR Results

Table 5 displays the overwhelming superiority of the Non-Ricardian alternative in each of the four pairs of regressions that were tested. It remains to consider the plausibility of the resulting parameter estimates. Table 6 gives the distribution of income and price elasticities for the two most complete specifications, Models 3 and 4. We note that there are no cases of extremely high absolute values of the estimated elasticities, and only a very modest representation of wrong signs, which led us to forgo the daunting task of “imposing” zero coefficients within the framework of SUR estimations.

Conclusions

We believe that the exercises reported here add weight to the growing body of evidence against the proposition that a strict interpretation of Ricardian Equivalence improves on our profession’s traditional use of a standard disposable income concept in explaining consumer behavior. This result holds both when the two competing income variables are for the current year only, and when they are constructed as four-year averages with declining weights for longer lags. We feel, too, that we have opened up a new area for testing Ricardian equivalence, in dealing with a single type of consumption (in our case residential electricity demand) rather than simply the aggregate of all types. This has enabled us to explore a rich body of data, for which estimates of income, but not aggregate consumption, were available. The field is obviously open for further testing along these lines, using whatever other components of consumption might be found to be reliably available at the state level.

TABLE 1
Results of Binomial Comparisons

A. Electricity Demand a Function of Income and Relative Price

1. Income Variables Both “Current”

Non-Ricardian Wins	Ricardian Wins	Ties	Ties Defined by $\left R_d^2 - R_r^2 \right $ less than
44	1	0	---
44	0	1	.001
44	0	1	.002
43	0	2	.003
43	0	2	.004
43	0	2	.005
42	0	3	.010

(Excluded States: MD, HI, NV, OR, TN)

2. Income Variables Both “Expected”

42	3	0	---
42	3	0	.001
42	3	0	.002
40	3	2	.003
40	3	2	.004
39	3	3	.005
36	3	6	.010

(Excluded States: MD, HI, OR, TN, WY)

Table 1 (cont.)

Electricity Demand a Function of Income, Relative Own Price,
Relative Gas Price, Weather

Non-Ricardian Wins	Ricardian Wins	Ties	Ties Defined by $ R_d^2 - R_r^2 $ less than
3. <u>Income Variables Both "Current"</u>			
38	10	0	---
33	7	8	.001
29	6	13	.002
27	6	15	.003
26	5	17	.004
25	5	18	.005
18	2	28	.010

(Excluded States: MD, WY)

4. Income Variables Both "Expected"

35	13	0	---
29	10	9	.001
26	10	12	.002
23	9	16	.003
20	8	20	.004
18	6	24	.005
12	4	32	.010

(Excluded States: MD, WY)

TABLE 2**Odds Favoring Non-Ricardian Over Ricardian Income Definition**

Regressions Being Compared	Odds	Probability Levels	
		Non-Ricardian	Ricardian
1a vs 1b	44,690/1	.99998	.00002
2a vs 2b	3,302/1	.9997	.0003
3a vs 3b	32/1	.9690	.0310
4a vs 4b	6.75/1	.8518	.1482

TABLE 3**Quartiles of Income Elasticity Estimates**

	<u>1st</u> <u>Quartile</u>		<u>Median</u>		<u>3rd</u> <u>Quartile</u>
1a +	.471443	+	.697906	+	1.027270
1b +	.469007	+	.692986	+	0.975978
2a +	.449240	+	.695293	+	1.01148
2b +	.25	+	.716682	+	1.099808
3a +	.25	+	.360384	+	0.535428
3b +	.25	+	.25	+	0.512024
4a +	.25	+	.345597	+	0.549010
5b +	.25	+	.267092	+	0.5232416

TABLE 4**Quartiles of Own-Price Elasticity Estimates**

	<u>1st</u> <u>Quartile</u>		<u>Median</u>		<u>3rd</u> <u>Quartile</u>
1a	0	-	.052776	-	.273266
1b	0	-	.026310	-	.378010
2a	0		0	-	.223671
2b	0		0	-	.316261
3a	0	-	.037986	-	.266045
3b	0	-	.081454	-	.321324
4a	0	-	.115089	-	.292092
4b	0	-	.107232	-	.321216

TABLE 5

Odds Favoring Non-Ricardian Over Ricardian Income Definition

<u>Regressions Being Compared</u>	<u>Bayesian Posterior Odds</u>	<u>Log- Likelihood Test (df=1)</u>	<u>Probability Non-Ricardian Is Preferred</u>
1a vs 1b	>10,000/1	289.19	>.9999
2a vs 2b	>10,000/1	216.12	>.9999
3a vs 3b	>10,000/1	17.69	>.9999
4a vs 4b	126.6/1	10.95	.9922

TABLE 6
Distributions of Income and Price Elasticity
Estimates Using SUR Procedures

<u>Estimated Income Elasticities</u>						
<u>Model</u>	<u>< 0</u>	<u>0 to 0.25</u>	<u>0.25 to 0.50</u>	<u>0.50 to 0.75</u>	<u>0.75 to 1.00</u>	<u>1.00 to 1.25</u>
3a	3	13	8	6	0	1
3b	3	13	8	6	0	1
4a	5	8	11	5	1	1
4b	5	9	11	4	1	1
<u>Estimated Own-Price Elasticities</u>						
<u>Model</u>	<u>>0</u>	<u>0 to -0.25</u>	<u>-0.25 to -0.50</u>	<u>-0.50 to -0.75</u>		
3a	1	15	13	2		
3b	1	15	13	2		
4a	1	15	10	5		
4b	2	14	12	3		

APPENDIX I

Data Sources

This paper has relied on a new and large data set which contains a variety of information which may be useful for testing various macroeconomic hypotheses with subnational data. It includes, but is not limited to, data on personal income, population, debt, electricity consumption, temperature fluctuations, and taxes.

Q -- the residential electricity consumption, is millions of kilowatt hours of electricity consumed by residential customers, divided by state population to give per capita consumption.

Pe -- the price of electricity is the total revenue, in thousands of dollars, earned by utility companies from their residential customers of electricity, adjusted for inflation using the GDP deflator, divided by millions of kilowatt hours of electricity consumed by residential customers. All electricity data are from the Statistical Year Book of the Electric Utility Industry, 1963-1994.

Pg -- the price of natural gas is calculated by taking the nominal price of 1,000 cubic feet at standard temperature and pressure of natural gas to residential customers. This is converted to the inflation-adjusted price of natural gas, relative to the real price of electricity to residential customers. Data are from the Natural Gas Annual, 1990 for 1967-1990 data, the Natural Gas Annual, 1994 (microfiche from Northwestern University Library), for 1991-1994, and by estimating that state residential prices for 1963-1966 were in the same proportions to total prices per 1,000 cubic feet nationally as was the case for 1967. Data for Hawaii for 1963-1979 were estimated by assuming that the ratio of Hawaii's price to that in California was the same as the average ratio, 3.23, for Hawaii/California during 1980-1991.

Yd -- personal disposable income per capita, is from the Survey of Current Business, July issues, various years. It is deflated by the GDP deflator.

Yr -- Ricardian disposable income per capita, adjusts Yd by subtracting net new issues of state debt, adjusted for inflation and expressed in per capita terms. Data are from the Commerce Department series Governmental Finances in 19xx and from State and Local

Government Finances in 19xx, various years. For the four year moving averages, data are used from 1959-1994. The increment to federal government debt is apportioned among the states on the basis of their share of total payments of federal personal income tax. Data are from the U.S. Treasury Department, Statistics of Income.

Weather variables -- Heating degree days are calculated by taking daily maximum and minimum temperatures, in degrees Fahrenheit, for a reporting station in a state, averaging these temperatures, and subtracting the result from 65 degrees. If the average temperature is greater than 65 degrees, the heating degree day number is 0 for that day. The cooling degree day is calculated by taking the average temperature, as above, and subtracting 65 degrees from it. Data are from the Climatological Atlas of the U.S. National Weather Service. A reporting station close to the state's population center was selected for each state.

APPENDIX II

Review of the Ricardian Equivalence Literature

The Ricardian literature began with Barro's (1974) article, which discussed several of the objections to the hypothesis, such as finite lives. The literature on Ricardian equivalence is immense, and many articles have looked at consumption, with the majority rejecting the hypothesis.

An implication of the Ricardian equivalence hypothesis is that people should not be affected in their consumption by whether government spending is financed by debt or taxes. People should "pierce the veil of debt," and be unaffected in spending by the mix of debt and taxes.

Perhaps the strongest empirical support for the hypothesis came from Kormendi's (1983) comparison of the effects of government spending and taxes. Kormendi found that changes in government spending depress consumer spending, while changes in taxes had no effect on consumer spending. However, Feldstein and Elmendorf (1987) countered that Kormendi's results are sensitive to World War II data, when shortages, rationing and patriotic appeals were operating at the same time as a deficit-financed increase in government spending. When the war years are excluded, Kormendi's results are reversed.

Evans (1993) uses data from 19 countries to test Ricardian equivalence and a 1985 model by Blanchard. Ricardian equivalence is "resoundingly rejected. Moreover, the estimated deviation from Ricardian equivalence is roughly what one would expect if households faced perfect insurance markets and did not have altruistic bequest motives."

Bagliano (1994), using 1960-90 data from the United Kingdom on pre-announced changes in income taxes, rejects Ricardian equivalence because consumption reacts strongly to

fiscally induced changes in current disposable income. The effect comes from the semi-durable and durable components of spending.

Barsky et al. (1986) give a model of consumption with an MPC closer to a non-Ricardian one which ignores future tax liabilities.

Fuster-Perez (1993) tests the Ricardian hypothesis against the non-Ricardian consumption model for Spain, Italy, Germany, France, and the UK. He finds that consumers do "take into account the public finance decisions."

Bailey (1993) notes that Ricardian equivalence may hold if, for instance, property taxes are capitalized into property values, and if bequests are an argument in the utility functions of parents.

Dalamagas (1993) separated the effects of debt illusion for countries ranked by their debt/GDP ratios. He found that debt illusion (government bonds perceived at least partially as net wealth) held for solvent (low debt/GDP ratio or net creditors) countries, while Ricardian equivalence was lent credibility by the group of high debt/GDP countries. In another article (1992), Dalamagas apparently did the same work as in the 1993 paper. The 1992 article had data for 52 countries.

Trostel (1993) in a theoretical paper finds that non-lump-sum taxes can result in deficits with increased consumption effects, even if bonds are not net wealth. This is because of the distortionary effects of taxes.

Seater (1993) provides a review article on the Ricardian literature. Seater remarks that Ricardian equivalence does not hold strictly, but that despite contradictory empirical results, it should not be ruled out as an approximation. The lack of support for crowding out--the effect of debt changes on interest rates--supports Ricardian equivalence.

Graham (1992) cautions that selection of the consumption variable affects the outcome of Ricardian tests.

Gruen (1991) gives a humorous test of Ricardian equivalence. Over six hundred Australian economics students were asked to estimate the level of outstanding Australian federal government debt. Eleven academic economists were asked to predict the proportion of students who had a "rough idea" of the debt level. Gruen found that the students' knowledge was very meager, and that the academics had overestimated the proportion of students with a "rough idea" of the debt by five fold. We can infer that debt cannot matter much if nobody knows what it is.

Cadsby and Frank (1991) tested Ricardian equivalence in the laboratory with an overlapping generations model. Observed deviations from the Ricardian model were small.

Kotlikoff, Razin, and Rosenthal (1990) describe a model in which Ricardian equivalence does not hold even with altruistic bequest motives. This comes from strategic bargaining between parents and children over bequests and children's behavior (the Rotten Kid Theorem?).

Thornton (1990) gives evidence for 16 OECD countries over 1975-86 which supports Ricardian equivalence.

Feldstein (1990) finds that with income uncertainty Ricardian equivalence will not hold.

Haque and Montiel (1989) find that borrowing constraints in developing countries cause Ricardian equivalence to fail.

Bibliography

- Bagliano, Fabio C.**, "Do Anticipated Tax Changes Matter? Further Evidence from the United Kingdom," Ricerche Economiche, 48:2, June 1994, 87-108.
- Bailey, Martin J.**, "Note on Ricardian Equivalence," Journal of Public Economics, 51:3, July 1993, 437-46.
- Barro, Robert J.**, "Are Government Bonds Net Wealth?," Journal of Political Economy, 82:6, November-December 1974, 1095-1117.
- , "The Ricardian Approach to Budget Deficits," Journal of Economic Perspectives, 3:2, Spring 1989, 37-54.
- Barsky, Robert B., Gregory N. Mankiw, and Stephen P. Zeldes**, "Ricardian Consumers with Keynesian Propensities," American Economic Review, 76:4, September 1986, 676-91.
- Bernheim, B. Douglas**, "Ricardian Equivalence: An Evaluation of Theory and Evidence," NBER Macroeconomics Annual: 1987 (Cambridge, MA and London, 1987), 263-304.
- Bomberger, William A.**, "The Effects of Fiscal Policies When Incomes Are Uncertain: A Contradiction to Ricardian Equivalence: Comment," American Economic Review, 80:1, March 1990, 309-12.
- Box, G.E.P. and G.C. Tiao**, Bayesian Inference in Statistical Analysis (Reading, MA: 1973).
- Breusch, T.S. and A.R. Pagan**, "The Lagrange Multiplier Test and Its Applications to Model Specifications in Econometrics," Review of Economic Studies, 47:2, 1980, 239-53.
- Cadsby, Charles Bram and Murray Frank**, "Experimental tests of Ricardian Equivalence," Economic Inquiry, 29:4, October 1991, 645-64.

- Dalamagas, Basil A.**, "How Rival are the Ricardian Equivalence Proposition and the Fiscal Policy Potency View?," Scottish Journal of Political Economy, 39:4, November 1992, 457-76.
- , "Fiscal Illusion and the Level of Indebtedness: An International Comparison," South African Journal of Economics, 61:1, March 1993, 45-58.
- Evans, Paul**, "Consumers Are Not Ricardian: Evidence from Nineteen Countries," Economic Inquiry, 31:4, October 1993, 534-48.
- Feldstein, Martin and Douglas W. Elmendorf**, "Taxes, Budget Deficits and Consumer Spending: Some New Evidence," NBER Working Paper #2355, August 1987, n.p.
- Feldstein, Martin**, "The Effects of Fiscal Policies When Incomes Are Uncertain: A Contradiction to Ricardian Equivalence," American Economic Review, 78:1, March 1988, 14-23.
- Flavin, Marjorie**, "Ricardian Equivalence: An Evaluation of Theory and Evidence: Comment," NBER Macroeconomics Annual: 1987 (Cambridge, MA and London, 1987), 304-09.
- Fuster-Perez, M. Luisa**, "La Hipotesis de equivalencia ricardiana: Un analisis empirico en los paises de la comunidad europea," Investigaciones Economicas, 17:3, September 1993, 495-506.
- Graham, Fred C.**, "On the Importance of the Measurement of Consumption in Tests of Ricardian Equivalence," Economics Letters, 38:4, April 1992, 431-34.
- Gruen, David W.** "What People Know and What Economists Think They Know: Surveys on Ricardian Equivalence," Australian Economic Papers, 30:56, June 1991, 1-9.

- Haque, Nadeem U. and Peter Montiel**, "Consumption in Developing Countries: Tests for Liquidity Constraints and Finite Horizons," Review of Economics and Statistics, 71:3, August 1989, 408-15.
- Kormendi, Roger C.**, "Government Debt, Government Spending, and Private Sector behavior," American Economic Review, 73:5, December 1983, 994-1010.
- Kotlikoff, Laurence J., Assaf Razin, and Robert W. Rosenthal**, "A Strategic Altruism Model in Which Ricardian Equivalence Does Not Hold," Economic Journal, 100:403, December 1990, 1261-68.
- Leiderman, Leonardo and Mario I. Blejer**, "Modeling and Testing Ricardian Equivalence: A Survey," International Monetary Fund Staff Papers, 35:1, March 1988, 1-35.
- Perasso, Giancarlo**, "The Ricardian Equivalence Theorem and the Consumption Function: A Survey of the Literature," Rivista Internazionale di Scienze Economiche e Commerciali, 34:7, July 1987, 649-74.
- Plosser, Charles I.**, "Ricardian Equivalence: An Evaluation of Theory and Evidence: Comment," NBER Macroeconomics Annual: 1987 (Cambridge, MA and London, 1987), 309-13.
- Savage, Linus**, "The Subjective Basis of Statistical Practice," manuscript written at the University of Michigan Department of Statistics.
- Savin, N.E. and R.J. White**, "The Durbin-Watson Test for Serial Correlation With Extreme Small Samples and Many Regressions," Econometrica, 45:6, November 1977, 1984-96.
- Seater, John J.**, "Ricardian Equivalence," Journal of Economic Literature, 31:1, March 1993, 142-90.

Thornton, Daniel L., "Do Government Deficits Matter?," Federal Reserve Bank of St. Louis Review, 72:5, Sep-Oct 1990, 25-39.

Trostel, Philip A., "The Nonequivalence between Deficits and Distortionary Taxation," Journal of Monetary Economics, 31:2, April 1993, 207-27.

White, H., "A Heteroscedasticity Consistent Covariance Matrix Estimator and a Direct Test of Heteroscedasticity," Econometrica, 48:4, July 1980, 817-18.

Zellner, Arnold, "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," Journal of the American Statistical Association, 57:298, June 1962, 348-68.

Zellner, Arnold and Chung-ki Min, "Bayesian Analysis, Model Selection and Prediction," in W.T. Grandy, Jr. and P.W. Milonni (eds), Physics and Probability: Essays in Honor of Edwin T. Jaynes (Cambridge: 1993), pp. 195-206.

ENDNOTES

¹This paper has had a long gestation period, during which we have received comments from many colleagues, mostly at the University of Chicago, UCLA and the Catholic University of Chile. In addition to expressing our general appreciation for these comments, we would like to give special thanks to Arnold Zellner, whose comments were particularly detailed and generous. Of course we alone are responsible for any deficiencies that may remain.

²We would consider it a quite natural corollary of this view that households in general leave to future generations somewhat more than what could be justified by a pure "bequest motive". In addition to what we want to leave our heirs (which might even be zero in some cases), there is the wealth that we have held to protect ourselves against future contingencies (which is likely to be positive at the moment of our demise even when the desired bequest is zero).

³The experiments reported in this section permit four comparisons as between current and permanent income concepts. As each of these covers 48 or 49 states (in 2 models Hawaii was excluded), there are 194 individual comparisons covering pairs of regressions in which the only difference is the substitution of a permanent for a current disposable income variable. Of these 194 comparisons, current disposable income generates a lower sum of squared residuals in 84 instances, and permanent disposable income in 110 cases.