

ARTHUR M. OKUN\*

*Brookings Institution*

# *The Personal Tax Surcharge and Consumer Demand, 1968–70*

THE REVENUE AND EXPENDITURE CONTROL ACT of 1968 was the key measure responsible for a large fiscal swing toward restraint from a full employment deficit of roughly \$10 billion in the first half of 1968 to a full employment surplus of about the same size in the first half of 1969.<sup>1</sup> That fiscal restraint was intended to halt the boom and thereby to curb inflation. As everyone knows, it did not stop inflation—nor has any of the host of other economic policy actions undertaken in the United States since 1967. The linkage between halting a boom and curbing inflation has turned out to be a much more difficult and less reliable process than the profession judged three years ago.

An even more serious criticism of the 1968 act is that it did not halt the boom as promptly or decisively as had been predicted by its proponents. Actual output exceeded estimated potential for a year after the tax surcharge and the expenditure hold-downs were instituted. Yet the record does

\* Richard Mullins assisted in the research. This paper was made possible by the generous contribution of time (human and computer) and effort by Otto Eckstein, Charles Warden, and Fred Bamber of Data Resources, Inc.; Harold Shapiro of the University of Michigan; Maurice Liebenberg and George Green of the U.S. Department of Commerce, Office of Business Economics; and Lawrence Klein of the University of Pennsylvania. Constructive comments on an earlier draft were received from George Jaszi and Robert Eisner, as well as from several participants in the Brookings Panel on Economic Activity.

1. Nancy H. Teeters, "Budgetary Outlook at Mid-Year 1970," *Brookings Papers on Economic Activity* (2:1970), p. 304; and *Economic Report of the President together with the Annual Report of the Council of Economic Advisers, February 1971*, pp. 24 and 73.

show a marked deceleration in the pace of economic activity starting in mid-1968. Whereas real GNP had grown at a 6½ percent annual rate during the first two quarters of 1968, it decelerated progressively to a rate of 3½ percent in the latter two quarters of 1968 and 2½ percent in the first two quarters of 1969.

Thus the fiscal restraint program failed to achieve its objective with respect to prices and did not even fully realize the desired prompt slowdown of aggregate activity. In this sense, the surcharge period offers a dramatic contrast to the experience that followed the tax cut of 1964, when economic activity accelerated markedly, just as the proponents had promised. Discussing that earlier experience in a paper completed in September 1965, I began with the statement, "The best known fact about the Revenue Act of 1964 is that, in the year and a half since it took effect, economic activity has expanded briskly."<sup>2</sup> Basically, I am obliged to begin the current discussion of the Revenue Act of 1968 by pointing out that the best-known fact about it is that inflation has continued unchecked since its enactment.

The second sentence of my paper on the 1964 tax cut, however, said: "But such *post hoc, propter hoc* reasoning will never do." That statement remains true in 1971. The fact that the fiscal restraint program of 1968 didn't do all it was supposed to do doesn't prove it did nothing. To what extent was it responsible for the significant deceleration in activity that followed it? To what extent did it curb demand *ceteris paribus*, only to be swamped by a strengthening of private demand that would have occurred quite independently of the fiscal restraint? To what extent did it have restraining effects that were offset by a shift toward ease in monetary policy?

These questions deserve answers.<sup>3</sup> In this paper I will focus on only a single aspect of these questions by reviewing one element of the program of fiscal restraint—the 10 percent surcharge on personal income taxes. To estimate the extent to which consumption was directly curbed by the added personal income taxes, I shall first describe the nature of the change in the tax law; next discuss some theoretical positions that might generate different expectations about the probable effectiveness of the surcharge; and finally examine the empirical evidence on consumer demand in 1968–70, as

2. Arthur M. Okun, "Measuring the Impact of the 1964 Tax Reduction," in Walter W. Heller (ed.), *Perspectives on Economic Growth* (Random House, 1968), p. 28.

3. Interesting answers are suggested by Lawrence R. Klein in "An Econometric Analysis of the Revenue and Expenditure Control Act of 1968–69" (paper prepared as a contribution to the forthcoming Musgrave Festschrift).

interpreted by the consumption equations of four well-known econometric models.

### **The Nature of the Surcharge**

The Revenue and Expenditure Control Act of 1968 was signed into law by the President on June 28. Its enactment was the culmination of a bitter legislative battle that lasted for more than ten months after President Johnson proposed a 10 percent surcharge on corporate and individual income taxes on August 3, 1967.

The act provided for a surcharge of 10 percent on federal corporate income taxes, retroactive to January 1, 1968. As it was subsequently extended, the corporate surcharge continued at the 10 percent rate through calendar year 1969; it then dropped to a 5 percent rate for the first half of 1970, statutorily translated into a 2½ percent rate on profits for calendar year 1970.

The limitations on federal budget outlays contained in the act called for expenditures to be reduced by \$6 billion below levels set forth in the January 1968 budget document for fiscal year 1969, but the act exempted outlays on Vietnam, interest on the public debt, veterans' benefits and services, and social security benefits.

The personal tax surcharge, which is the focus of this paper, was enacted at a 10 percent rate, retroactive to April 1, 1968 and continuing to June 30, 1969. At an annual rate, it added about \$8 billion to aggregate liabilities for personal income taxes. The law exempted low-income taxpayers, such as families of four with incomes of less than \$5,000; but this relief provision amounted to only \$300 million per year in the aggregate. For a typical family of four with a \$10,000 annual income, the surcharge added \$111 to tax liabilities on a full-year basis; at a \$25,000 income, it was worth \$441.

Specifically, the 1968 law provided for a 7½ percent increase in the liabilities of covered taxpayers for the calendar year 1968 (reflecting a 10 percent surcharge for three-quarters of the year), and 5 percent for 1969 (a 10 percent surcharge for half the year). The withholding rate of personal income taxes was raised on July 15, 1968, to take account of the surcharge. Because of its retroactive feature, the surcharge provision reduced refunds and increased final tax payments on 1968 liabilities in the spring of 1969. Subsequently, in 1969, the surcharge was extended at a 10 percent rate for the second half of the calendar year, and at a 5 percent rate for the first half

of calendar 1970. It expired in effect on June 30, 1970, although the law provided for a 2½ percent increase in liabilities for the 1970 calendar year.

The payment of federal income taxes is recorded in the national income accounts at the time the taxpayer actually parts with cash, either through withholding or through payment of estimated or final installments. From the liability rates and the withholding rates that prevailed before and after the enactment of the surcharge, the incremental federal personal income taxes attributable to the surcharge (for given before-tax incomes) can be estimated in a fairly straightforward manner. The resulting estimates, recorded in Table 1, were provided by Charles Waite of the Department of Commerce.<sup>4</sup> The series begins in 1968:3 and then takes a temporary jump in the first half of 1969, reflecting the retroactive feature; it steps down at the beginning of 1970 when the surcharge rate fell to 5 percent; and it finally drops to a negligible figure after mid-1970.

### **The Link to Consumption**

The question of how the consumer responded to the income tax surcharge merely asks to what extent the resulting cutback in disposable income held down consumption outlays and to what extent it held down personal saving. An arithmetic identity ensures that any change in income must be allocated between a change in consumption and a change in saving.<sup>5</sup> According to the historical record of aggregate quarterly data on household income and its disposition, most changes in disposable income are primarily reflected in altered consumption, at least after a lag of a few quarters. On the margin as well as on the average, the largest portion of disposable income is consumed. The record provides evidence on how much and how fast a typical change in disposable income is translated into a change in consumption. The same historical record, however, warns us that disposable income is not the only determinant of consumption outlays. Such variables as stocks of consumer goods, wealth, liquid assets, and credit con-

4. The estimates are expressed in terms of seasonally adjusted annual rates. Seasonal adjustment of the nonwithheld portion is not a straightforward matter, unfortunately.

5. It seems safe to ignore the possibility that the tax increase led to a quantitatively significant change in the remaining minor component of disposable income, personal transfer payments; these consist of payments of interest on personal debt to businesses, and gifts to persons overseas. Personal transfers amounted to 2.6 percent of disposable income throughout the period.

**Table 1. Impact of Tax Surcharge on Federal Personal Income Tax Payments, 1968-70**

Billions of dollars, seasonally adjusted annual rates

<i>Year and quarter</i>	<i>Current dollars</i>			<i>1958 dollars<sup>a</sup></i>
	<i>Withheld</i>	<i>Nonwithheld</i>	<i>Total</i>	<i>Total</i>
1968:3	5.3	0.8	6.1	5.1
4	6.3	0.8	7.1	5.9
1969:1	6.5	4.2	10.7	8.8
2	6.7	4.2	10.9	8.9
3	6.9	0.2	7.1	5.7
4	7.1	0.2	7.3	5.8
1970:1	3.6	1.4	5.0	3.9
2	3.6	1.4	5.0	3.9
3	0.0	0.4	0.4	0.3

Source: Tabulation by Charles Waite, U.S. Department of Commerce, Office of Business Economics, December 23, 1970.

a. Deflated by implicit deflator for personal consumption expenditures.

ditions have been found (by at least some investigators) to influence consumption significantly. And even with all the determining factors that can be quantified, economists cannot explain perfectly the allocation of income between consumption and saving. The behavior of the consumer does not precisely follow statistical predictions and at times departs widely from such relationships.

Ever since economists have become interested in fiscal policy, they have operated generally on the fundamental premise that changes in after-tax income resulting from a change in personal tax rates are basically equivalent in their influence on consumption to changes in income arising from other sources. In fact, there is not much direct empirical evidence to validate (or refute) this fundamental premise. Following the big tax cut in 1964, consumer outlays corresponded very well with the assumption that the tax cut would affect consumer spending about as much and as rapidly as any other addition of equal size to spendable income. But most of the support for the fundamental premise is analytical. So long as a tax increase has no special effects on consumer expectations, asset values, or income distribution, and so long as it does not result in additional public services that were previously purchased in the private market, a tax increase of one dollar "ought" to offset a before-tax income gain of one dollar. In a sense, the premise is an implication of the more fundamental postulate that consumers behave ra-

tionally with respect to their budget constraints. After-tax income sets the constraint on the sum of consumption and saving; it is the income available for allocation by the household. If, by coincidence, a family were to experience a sudden but sustained rise in income before taxes at the same time and by the same amount that tax rates were increased, the household would be neither better off nor worse off; there would be no clear reason to expect it to raise or lower its standard of living (or thus to alter its rate of saving). A family's spending is not expected to depend significantly upon whether it is paid in check or in cash, or whether payday is Friday or Monday; by the same reasoning, its consumption is not expected to depend merely on the way its disposable income is made up of before-tax income and subtractions for income taxes.

#### RESERVATIONS AND QUALIFICATIONS

Although the fundamental premise is intuitively appealing, it is merely suggestive, rather than conclusive. A tax change could have special effects on expectations and on other subjective determinants of consumption. In any particular episode, it is easy to offer reasons that might make the consumer respond differently and uniquely. In the case of the surcharge, several such reasons have been suggested. The most important of these hypotheses is related to the permanent income view of consumer demand; it will be explored in detail below. But first, several other hypotheses deserve a brief analytical review.

*The triviality argument.* Many laymen seem puzzled that economists ever expect a significant effect on consumer expenditures from a change in after-tax income that amounts to only a dollar or two a week for middle-income families. Why, they ask, should that change anybody's standard of living? It seems too trivial to matter.

Such arguments have never carried much weight within the profession, however. Although the economist can imagine income changes that may be too trivial to matter, he can't decide whether they would be too trivial to affect consumption or too trivial to affect saving, and they have to affect at least one of the two. The rhetorical question can be turned on its head: Why should a small change in income affect people's decisions on the desirable rate of increasing their wealth? The implication of that question is that any small increment (or decrement) in income will be matched by a change in consumption and have no effect on saving.

Moreover, discontinuities at the household level could still be consistent with continuity in the behavior of aggregates. Even if every family were prepared to hold its consumption constant unless and until its income changed enough to cross some threshold, some families would be close to their threshold at any one time and could be pushed into a substantial change of consumption by a tiny change in income. Perhaps it is merely a professional bias, but the economic analyst relies on continuity of aggregate responses to all sorts of small changes in the economy, and rarely does he find empirical evidence that contradicts this assumption.

*The ratchet argument.* Nearly a generation ago, James Duesenberry and Franco Modigliani provided good analytical and empirical evidence of an asymmetrical response of consumption to increases and decreases in income. Generally, consumption declines less when income falls below a previous peak than it rises in response to an equal increase in income to a new peak.<sup>6</sup> There is a ratchet effect and a resistance to belt-tightening when income declines.

But the ratchet argument seems to have little relevance to the tax surcharge experience. Of course, most American workers experienced a slight dip, of roughly 1 percent, in their take-home pay at the moment the surcharge withholding rates took effect. But, for most families, changes in before-tax income must have swamped this effect within a short interval. From 1968:2 to 1968:3, per capita disposable income actually rose in real as well as nominal terms. The ratchet argument would have important implications for aggregate consumption only if an extraordinary fraction of American families experienced income declines or remained below their previous peak incomes for a sustained period. And there is no evidence to suggest that such was the case.

The variability of income for individual households is enormous. As an earlier study showed, seven-eighths of all families reported that some event between January 1964 and May 1965 had changed their income by 10 percent or more. Two-thirds reported that some event had *decreased* their income by such an amount during this period, although most of these families also experienced events resulting in even larger income increases.

6. James S. Duesenberry, *Income, Saving, and the Theory of Consumer Behavior* (Harvard University Press, 1949); and Franco Modigliani, "Fluctuations in the Saving-Income Ratio: A Problem in Economic Forecasting," in *Studies in Income and Wealth*, Vol. 11, *Conference on Research in Income and Wealth* (National Bureau of Economic Research, 1949).

Only about half of all families had incomes in 1964 that were within 10 percent of their 1963 incomes.<sup>7</sup> As these data suggest, only an infinitesimal fraction of families have exactly the same before-tax incomes two years in a row. It is hard to believe that a significant fraction of total income belonged to the group of families that was pushed from no change (or a minuscule rise) in income into the minus zone as a result of the surcharge. The ratchet argument applies to absolute declines in (real or money) income, and the surcharge itself was rarely responsible for absolute declines.

*Interaction with inflation and inflationary expectations.* Since the impact of recent or expected price increases on consumer demand is uncharted, it attracts conjecture and speculation. If consumption seems unusually weak while prices are rising strongly (as in 1968:2 or 1970:3), one tends to suspect that the consumer is resisting inflation or adjusting to a reduction in the real value of his liquid assets. On the other hand, when consumer demand seems particularly strong in an inflationary period (like 1966:1), one wonders whether inflationary expectations are encouraging consumers to buy now in order to beat subsequent price increases. And indeed both conjectures may conceivably be right.

Putting the tax surcharge into the picture permits endless variations, in either direction, on the causal arguments. To sample one line of conjecture: When people feel squeezed by inflation, they want to rationalize actions to maintain consumption and trim saving. The tax surcharge gave them an excuse for saving less, and they grabbed it, especially since they were fairly liquid and found it easy to borrow. Another conjecture invokes a self-denying prophecy: People felt uncertain because of the inflation, and they thought the government was stopping inflation when the tax surcharge was passed; therefore they might have spent more of their income and stimulated the inflation.

*Advance perception of the tax increase.* The long legislative debate over the tax surcharge gave the consumer an opportunity to anticipate its impact. He may accordingly have reduced consumption before the enactment of the surcharge, particularly on durable goods involving installment charges. Indeed, in advance, many people seemed to overestimate the impact of the tax surcharge on their incomes, even interpreting it as an added liability amounting to 10 percent of their *income* rather than 10 percent of

7. George Katona and Eva Mueller, *Consumer Response to Income Increases* (Brookings Institution, 1968), pp. 54-67.

their *taxes*. Thus, the anticipation might have been worse than the realization.

Such highly subjective, ad hoc conjectures defy systematic verification, but they cannot be dismissed.

### The Permanent Income Hypothesis

Among all the conceivable reasons that the surcharge could have had a significantly different impact from other cutbacks in the growth of disposable income, the most serious and most challenging is the permanent income hypothesis—the proposition that consumption is altered much less in response to changes in income that are viewed as transitory than in response to those viewed as permanent. As stated strongly by Robert Eisner:

The basic economic error of those who saw in the 10 percent income tax surcharge an adequate measure against inflation may be charged to failure to take into account the implications of the permanent income hypothesis . . . . Dealing in terms of equilibrium or permanent income, an increase in tax withdrawals of some ten or eleven billion dollars per year may be expected to have a very substantial depressant effect on aggregate demand. . . . But the tax increase legislated in the surcharge did not represent a corresponding reduction in *permanent* after-tax income.<sup>8</sup>

#### PERMANENT INCOME AND TEMPORARY TAX CHANGES

The analysis of a temporary tax change in light of the permanent income hypothesis raises a number of important theoretical and empirical issues.

Among the assumptions and implications of the permanent income hypothesis, the one that makes it relevant to the tax surcharge is the view that households gear their standards of living, not to measured income in a single calendar quarter or year, but to their longer-run income expectations or “permanent income.” When a family experiences a change in current income that it expects to be sustained through time, the response of consumption is full and reasonably prompt. However, when the change in cur-

8. Robert Eisner, “Fiscal and Monetary Policy Reconsidered,” *American Economic Review*, Vol. 59 (December 1969), p. 898. Eisner also argues that investment could not have been curbed to any significant degree by the temporary corporate surcharge, and I find his argument on that issue more persuasive.

rent income is viewed as transitory and does not alter expected future income, the household tends to spread the transitory or windfall component over a long horizon, allowing only a small portion to alter consumption in the current year.<sup>9</sup>

In its loosest qualitative terms, the hypothesis has overwhelming appeal. Anybody who does any saving reveals that his relevant time horizon for planning is longer than the current period. Unless wealth accumulation is his primary and ultimate goal, he must be concerned about fitting consumption to his needs over a longer time horizon. If a family gets substantial income receipts that it views as windfalls, the impact on current consumption is bound to be smaller than would be the case from an equal addition to income that is expected to be recurrent and permanent. The controversial issue is how much difference this makes: What types of income items do consumers identify as purely transitory and under what circumstances? What is the length of the horizon they focus on in their consumption decisions?

Even between the principal pioneers of the permanent income view, the judgments are far apart: Friedman uses a three-year horizon, assessing the marginal propensity to consume out of windfalls at roughly one-third, while Modigliani–Brumberg take a lifetime horizon, implying an annual impact as small as 0.04. These authors agree, however, in limiting the scope of the permanent income hypothesis to consumer expenditures on non-durables and services. They distinguish between the purchase of durable goods and the enjoyment of a flow of services from durables; both analyses view the acquisition of additional durable goods as one of the many ways in which transitory income may be added to wealth.

#### EMPIRICAL TESTS

The response of consumers to windfalls can be investigated directly only through the few surveys of household budgets in which sources of income have been identified in sufficient detail that presumptive windfalls can be isolated. Several tests have been performed by Ronald Bodkin and others

9. Milton Friedman, *A Theory of the Consumption Function* (Princeton University Press for the National Bureau of Economic Research, 1957); and Franco Modigliani and Richard Brumberg, "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data," in Kenneth K. Kurihara (ed.), *Post Keynesian Economics* (Rutgers University Press, 1954).

on a 1950 U.S. survey that isolated windfall income consisting largely of National Service Life Insurance dividends. Both in his initial report and in subsequent investigations, Bodkin found no evidence that the marginal propensity to consume was less from windfall income than from other income. Indeed, Bodkin ran into the problem that his data tended to prove too much: The marginal propensities to consume from windfall income generally exceeded those from regular income. This curious result lent some plausibility to Friedman's conjecture that the amount of the life insurance dividend might have been indirectly related to the permanent income of the household.<sup>10</sup>

Another group of studies investigated Israeli data in which the prime source of windfall income was restitution payments to German refugees. There the authors found a resounding difference between spending propensities, just as would be implied by the permanent income hypothesis. According to Mordechai Kreinin, the marginal propensity to consume out of windfalls was about one-fourth that out of other income, as revealed in the Israeli data.<sup>11</sup>

A persuasive reconciliation of the Bodkin and Kreinin findings was presented by Michael Lansberger.<sup>12</sup> The windfalls in the Kreinin study were large, of roughly the same order of magnitude as the regular income of the recipients. In contrast, the windfalls in the 1950 U.S. data averaged about 7 percent of the total income of recipients. Lansberger showed further that the marginal propensity to consume out of windfalls was, in the Israeli data, very strongly negatively related to the size of the windfall in relation to other income. This pattern suggested that households segregate large windfalls from regular income and treat them specially as capital account items, but that families typically lump small windfalls with regular income. The

10. Ronald Bodkin, "Windfall Income and Consumption," *American Economic Review*, Vol. 49 (September 1959), pp. 602-14; Margaret G. Reid, "Consumption, Savings and Windfall Gains," *American Economic Review*, Vol. 52 (September 1962), pp. 728-37; Roger C. Bird and Ronald G. Bodkin, "The National Service Life-Insurance Dividend of 1950 and Consumption: A Further Test of the 'Strict' Permanent-Income Hypothesis," *Journal of Political Economy*, Vol. 73 (October 1965), pp. 499-515; and Milton Friedman, "Comments," in Irwin Friend and Robert Jones (eds.), *Proceedings of the Conference on Consumption and Saving* (University of Pennsylvania Press, 1960), Vol. 2, pp. 197-98.

11. Mordechai E. Kreinin, "Windfall Income and Consumption—Additional Evidence," *American Economic Review*, Vol. 51 (June 1961), pp. 388-90.

12. Michael Lansberger, "Windfall Income and Consumption: Comment," *American Economic Review*, Vol. 56 (June 1966), pp. 534-40.

same results were suggested by a survey of Colorado families analyzed by Conrad Doenges.<sup>13</sup>

Indeed, no empirical study suggests that a negative windfall as small as the surcharge's dent of 1 percent of disposable income would be treated any differently from a permanent reduction in income of the same size. In short, the windfall view of the surcharge gets no support from the direct empirical tests.

That view is also subject to other reservations. First, influenced in part by the history of the Korean war "temporary" taxation, American citizens typically were skeptical that the tax surcharge would actually expire in a short time. Secondly, the very nature of the wartime inflationary economy of 1966–68 must have created enormous uncertainties about long-run prospects for both before-tax and after-tax real incomes. Some workers in defense plants must have doubted the viability of long-term careers in their Vietnam-related jobs. Other manufacturing workers who had unusual opportunities to work overtime must have questioned whether the phenomenon was temporary or the hallmark of a new age of perennial full (or over-full) employment. Married women must have asked the same questions about the ready availability of jobs for them. The new experience with inflation after almost a decade of abeyance must have raised important questions about the longer-run horizon among both people who were hurt by price increases and those who reaped gains. In that environment, a great many elements may have created a gap between current income and permanent income, as perceived by consumers. Any and all of these could have influenced consumer demand. So could anything that altered the degree of confidence people felt about the long run, given the precautionary character of some saving. The surcharge may have been one of the special factors, but it hardly seems reasonable to believe that it was a key element in the transitory components of household incomes.

Just as a matter of autobiographical fact, I wondered and guessed about these issues in 1967–68; as I reported in 1969: "Viewing the surcharge in advance, I felt that some discount on its restraining effect should be taken for its temporary character. There was no certainty about the right size of the discount, but 20 or 25 percent seemed to be in the ballpark."<sup>14</sup>

13. R. Conrad Doenges, "Transitory Income Size and Savings," *Southern Economic Journal*, Vol. 33 (October 1966), pp. 258–63.

14. Arthur M. Okun, *The Political Economy of Prosperity* (Brookings Institution, 1970), p. 96.

The Tax Reform Act of 1969 raises another issue about the interpretation of temporary and permanent income tax rates. That act provides for a cut in personal taxes of \$9.5 billion a year, with the reductions scheduled to take effect in stages lasting until January 1973. Thus, Congress adopted lower permanent tax rates but essentially enacted a temporary measure delaying their effective date. On a logical interpretation of Modigliani–Brumberg and (to a somewhat lesser degree) Friedman, the reduction in permanent tax rates should have been a major determinant of consumer outlays in 1970 and should have generated an unusually high rate of consumption in relation to current income. Recent consumption-income patterns do not confirm this view, nor has any permanent income exponent followed the logic of his theory to argue this case. Of course, many consumers may not even be aware of these forthcoming changes in the tax base and rates; but that in itself is relevant to the issue. To take the long horizon view of consumer budgeting seriously, one must believe that families are careful forecasters of their own financial future and energetic collectors of relevant information. A man who does not read news reports on such matters as changes in tax rates and in social security benefits is simply not a Modigliani–Brumberg man.

### **The Nature of the Experiment**

Any theoretical position that dismisses or strongly discounts the consumption impact of the tax surcharge would point to higher levels of consumer spending during the surcharge period than those implied by the alternative view that the surcharge was effective in curbing consumption. According to the windfall view of the surcharge, the expected differences would be concentrated in outlays for nondurable goods and services. These differences permit empirical tests of the alternative views.

Statistical relationships that use officially measured disposable personal income as the income variable in explaining consumption implicitly count on the surcharge to be as effective as any other cutback in spendable income, with no discount necessary to allow for the transitory income effect or for any other reason. This may be called the *full-effect* view of the surcharge. Of course, the standard equations allow for lags in the response of consumption to income; any income change results at first in a relatively small change in estimated consumer spending, with the effects building up

over time. If, during the surcharge period, consumption equations based on the full-effect view predicted consumer outlays with errors that were not unusually large and that averaged close to zero, the full-effect view would be confirmed.

The full-effect view may be contrasted with a *zero-effect* view that the surcharge had no influence on consumption (that is, its total impact was on saving). According to the zero-effect view, the relevant concept of income for predicting consumer expenditures during the surcharge period should be disposable personal income (as officially measured) *plus* the amount of extra personal taxes imposed by the surcharge. If predictions of consumption during the surcharge period based on this adjusted income concept yielded satisfactory results with errors symmetrically distributed around zero, the zero-effect view would be confirmed.<sup>15</sup>

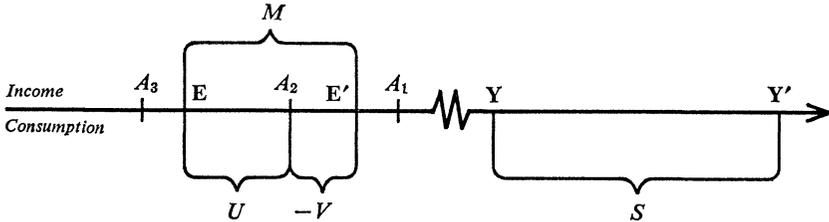
#### A SCHEMATIC PRESENTATION

These alternative hypotheses are shown schematically in Figure 1. The level of officially measured disposable income for any quarter in the surcharge period, consistent with the full-effect view, is designated by  $Y$ . The adjusted income concept based on the zero-effect view is  $Y'$ , which is higher than  $Y$  by  $S$ , the dollar amount of the surcharge. Now assume that there is an agreed set of "best" statistical relationships for estimating consumer demand, given the relevant concept of income. There will then be two alternative estimates of consumption, designated  $E$  and  $E'$ , based on the two income concepts,  $Y$  and  $Y'$ , respectively. Of course, because it assumes a higher value of relevant income,  $E'$  exceeds  $E$ . The distance between  $E$  and  $E'$  is the expected direct impact ( $M$ ) of the surcharge on consumer spending, as interpreted by the full-effect view.

Now suppose that the actual value of consumption for a given quarter is observed, and that it turns out to be  $A_2$  in the diagram. Then the prediction error based on the full-effect view is  $(A_2 - E)$ ; in accord with the convention that the error is measured as actual value minus predicted value, the underprediction yields a positive error designated as  $U$ . The prediction error based on the zero-effect view is  $(A_2 - E')$ , in this case a negative number designated as  $V$ . In general,  $M = U - V$ .

15. Neither the transitory-income view of the surcharge nor any of the other skeptical views cited above necessarily implies that the surcharge would have no effect whatsoever; the zero-effect view is simply a convenient benchmark of extreme skepticism.

**Figure 1. Schematic Diagram of Full-Effect and Zero-Effect Views of Surcharge**



The basic empirical test to discriminate between the full-effect and zero-effect hypotheses consists of comparing the size (in absolute terms, independent of sign) of the typical  $U$  and typical  $V$ . If  $A_2$  (between  $E$  and  $E'$ ) were the typical observation throughout the surcharge period, the results would point to a partial effect of the surcharge:  $U$  would be positive (reflecting underprediction by the full-effect view) and  $V$  would be negative (reflecting overprediction by the zero-effect view). In particular, if, as predicted,  $A_2$  is somewhat closer to  $E'$  than to  $E$ ,  $U$  is larger in absolute size than  $V$ , revealing the full-effect view as less accurate than the zero-effect view. In that event, however, an hypothesis of less than half of the full effect would be more accurate than either the zero- or full-effect view. If the typical consumption observation was  $A_1$  (greater than  $E'$ ),  $U$  would be larger (absolutely) than  $V$  and the zero-effect view would be the more accurate by far. However, in that case, with actual consumption higher than the estimate based on the zero-effect view, it would have to be inferred that some other stimulative factor was operative, even with the surcharge having zero effect. Similarly, if actual consumption ran consistently below  $E$ , as illustrated by  $A_3$ , then the full-effect view would be superior, but something besides the tax surcharge must have been working to hold consumption down.

Such an interpretation begins essentially with the hypothesis that the only forces consistently and significantly pushing consumption upward or downward during the surcharge period were the variables specified in the consumption equations and the surcharge. If actual consumption outlays remain, on average, between  $E$  and  $E'$ , this hypothesis is maintained. If, however, they lie outside that range, it becomes necessary to discard the hypothesis that other forces averaged out.

To the extent that the surcharge had any direct impact on consumption, it also would have had subsequent multiplier effects curbing the growth of before-tax income and accelerator effects dampening business investment. It might also have led to a change in monetary policy. None of these second-round or subsequent effects is reflected in  $M$ . That measure of direct impact simply records the first-round impact based on the full-effect view, when all other relevant variables are known, including the components of personal income, the relevant monetary and credit factors, the level of employment, and so on.

#### THE NEED FOR DYNAMIC SIMULATION

Since the empirical test focuses on the direct impact of the surcharge and not on multiplier effects, it is perfectly satisfactory to review consumer behavior during the surcharge period using *actual* values of disposable income (as interpreted by either the zero- or full-effect view), prices, money variables, and unemployment. A serious problem arises, however, because many of the consumption equations include as explanatory variables the past value of the consumption component that is being predicted. The proposition that expenditures for gasoline (for example) this quarter depend, in part, on previous expenditures for gasoline obviously does not imply a causal relationship. Rather, lagged expenditure is designed to reflect habit formation and persistence of consumer tastes through time. But the lagged terms also are useful for predictive purposes because they correct recent errors in the causal explanation of that component of consumption and because they may capture autocorrelated measurement errors in the data. Basically, in attributing substantial weight to past expenditures, the computer is telling the investigator: When consumption is higher than expected in one quarter, history teaches us that it is also likely to be somewhat higher than would otherwise be expected in the next quarter.

The lagged expenditure term thus permits predictions based on an incorrect theoretical view of consumption to be mechanically corrected and to approach the predictions based on the valid analytical view. It may become impossible to tell who is right and who is wrong if errors are mechanically corrected quarter after quarter. If the full-effect view of the surcharge is all wrong, the investigator may observe at the start of the surcharge period that consumption is higher than he anticipated; if he then uses these higher actual values to raise his predictions for subsequent quarters, he may fuzz

up the distinction between the zero-effect and full-effect views. Static simulation, the formal term for this quarter-by-quarter technique, will thus not produce efficient discrimination. As a result, one must turn to another technique known as dynamic simulation. In this experiment, the investigator is asked to "forecast" consumption for the surcharge period, knowing all the relevant determining variables *except* the actual values of consumption during that period. Whenever his statistical relationships say that current consumption depends in part on previous consumption, the investigator must rely on the values of previous consumption that he himself predicted.

The choice of the technique of dynamic simulation requires the further choice of a starting point. When should the experimenter be required to stop looking at the consumption data? There is good reason to blindfold him well before the surcharge period begins because the influence of the state of demand in the last known quarter diminishes over time; this practice makes the results during the surcharge period less sensitive to the particular conditions of demand in the last known quarter.<sup>16</sup> I selected 1967:2 as the last known quarter before seeing any of the simulations. I liked it because it was the quarter immediately preceding the surcharge proposal, because consumer expenditure was not behaving very abnormally in relation to income at that time, and because a head start of five quarters prior to the surcharge period is enough to dilute much of the influence of the jump-off point.

Thus, the experiment consists of "predicting" consumer spending during the surcharge period of 1968:3 to 1970:3 on the two extreme assumptions of full effect and zero effect. The *ex post* predictions are made with full information on actual economic activity prior to mid-1967 and on all relevant matters after mid-1967 *except* actual values of consumer expenditures.

#### THE MODELS

The consumption equations of four well-known econometric models were applied to this experiment: Data Resources, Inc. (DRI); Michigan's

16. The persistence of the effect depends on the coefficient of the lagged expenditure term. It decays exponentially through time at a rate depending on its distance from unity. A coefficient of 0.5 gets down to 0.06 of the initial error in four quarters, while one of 0.9 is still 0.66 at that point.

DHL-III; Office of Business Economics (OBE); and Wharton.<sup>17</sup> The consumption components in DRI are expressed in current dollars while those in the other three models are expressed in 1958 dollars; for purposes of comparison, all errors and predictions of each equation of the DRI model have been converted into 1958 dollars in the tables and discussion below.

The sample periods for the statistical estimates also differed among the models: They were 1954:1 to 1970:2 for DRI; 1954:3 to 1967:4 for Michigan; 1953:2 to 1968:4 for OBE; and 1953:3 to 1970:1 for Wharton. Ideally, the equations employed for this experiment should have used sample periods terminating before the onset of the surcharge in 1968:3. Since the official disposable income series used in the equations reflects the full-effect view of the surcharge, the inclusion of quarters after 1968:2 might tend to pull the estimated coefficients in the direction of that view. Only the Michigan model meets the criterion of purity; the OBE sample period is nearly pure. Wharton and DRI, on the other hand, include most of the surcharge period in their samples. Still, the surcharge period represented only about one-eighth of the total sample period, and none of the results below casts suspicion on this particular feature of the test.

The models vary in their degree of disaggregation of consumption expenditures, with three equations for Michigan; four for Wharton; five for OBE; and eleven for DRI. (A detailed listing of these equations is available on request from the author.) Nondurable goods and services outlays combined are explained in a single equation by the Michigan model. That relationship has three explanatory variables: the level of spending on nondurables and services in the preceding quarter (with a coefficient of 0.945), disposable income in the current quarter, and disposable income in the preceding quarter.

At the other extreme of disaggregation is the DRI model, which identifies

17. The most recent published descriptions of these models are Data Resources, Inc., *The Data Resources Econometric Forecasting System: A Preliminary Account* (DRI, November 1970); Saul H. Hymans and Harold T. Shapiro, "The DHL-III Quarterly Econometric Model of the U.S. Economy," in *The Outlook for 1970*, Seventeenth Annual Conference on the Economic Outlook at the University of Michigan, November 20-21, 1969 (1970); Maurice Liebenberg, Albert A. Hirsch, and Joel Popkin, "A Quarterly Econometric Model of the United States: A Progress Report," *Survey of Current Business*, Vol. 46 (May 1966), pp. 13-39; Michael K. Evans and Lawrence R. Klein, programmed by George R. Schink, *The Wharton Econometric Forecasting Model* (University of Pennsylvania, 1967). The OBE and Wharton models have been revised extensively, and new publications on both are forthcoming. In addition, some or all of the equations in the Data Resources and Michigan DHL-III models have been updated.

and explains eight separate components of nondurables and services. Each of these equations relies on disposable income for the current quarter as the single income variable. Most of the equations contain a relative price variable, and some also include population or a time trend. In addition, each includes a weighted average rate of consumption of the item in question over the preceding four quarters, whose predetermined weights range downward, counting from the preceding quarter, from 0.4 to 0.1. The coefficients on this lagged spending term are roughly 0.5 for most of the nondurable goods categories, but for gasoline and oil and for the services they run between 0.78 and 0.92. Coefficients of this magnitude ruled out a meaningful static simulation.

The OBE approach employs a single equation for nondurables and two for services. The nondurables equation includes average expenditures on nondurables for the preceding four quarters as one explanatory variable. It also has the special feature of distinguishing transfer payments from other components of disposable income; the estimated impact of transfer payments on consumption of nondurables is about one and one-half times as large as that of other disposable income, presumably reflecting the tendencies of transfer recipients, such as social insurance pensioners, to devote more of their incomes to nondurables and less to autos and saving. The two service components, housing and other services, are explained on a per capita basis, both with a term to correct for last period's error. The housing equation is unique in having no direct income variable; instead it uses the sum of expenditures on nondurables and services in the preceding four quarters, presumably as an indicator of how much is available for expenditure on housing.

The Wharton model devotes one equation each to nondurables and services: Current disposable income appears in both equations, as does the expenditure in question with a one-quarter lag. The coefficients on the two lagged expenditure terms differ widely: 0.98 for services and 0.37 for nondurables. In addition, the nondurables equation has a relative price term. A noteworthy "non-Keynesian" feature of these two Wharton equations is a money-stock term, which attributes to recent increases in broadly defined money (currency, demand deposits, and time deposits) a modestly stimulative influence on consumer outlays for nondurables and services.

Consumer expenditures on durable goods, excluding autos and parts, are explained in a single equation by three of the models. Michigan uses three variables: disposable income in the current quarter and in the im-

mediately preceding quarter, and expenditures on nonauto durables in the preceding quarter. The 0.87 coefficient on the lagged expenditure term shows strong persistence of demand patterns.<sup>18</sup> The OBE model accounts for nonauto durables with two explanatory variables: disposable income of the current quarter, and a liquid asset measure. In the Wharton model, the explanatory variables are disposable income of the current quarter, a relative price measure, the growth of the broadly defined money supply over the last four quarters, the unemployment rate and its change over the past four quarters, and residential construction activity.

In the DRI model, consumer expenditures on nonauto durables are divided into two categories with furniture and appliances separated from all other items. Disposable income and the unemployment rate of the current quarter are explanatory variables in both equations. In addition, the demand for furniture is taken to be positively related to recent housing starts and negatively related to current expenditures on automobiles and parts.<sup>19</sup>

Each of the models has an equation explaining consumer demand for automobiles. All include a "dummy" to allow for strike effects. In addition, DRI uses the following variables: current disposable income net of transfer payments, unemployment, the stock market, and an income change variable. In the Michigan model the explanatory variables are disposable income lagged one and two quarters; the unemployment rate for males of age twenty and over, lagged one and two quarters; a relative price variable; and lagged expenditures on automobiles of the preceding quarter. In OBE, the variables are disposable income net of transfers, the manufacturing work week, the stock of automobiles, relative prices, and a correction for the prediction error of the preceding quarter. Wharton's equation includes current disposable income net of transfers, a relative price variable, the unemployment rate, a consumer credit dummy, and a variable reflecting the interest rate structure. Both OBE and Wharton explain automobiles and parts net of expenditures on mobile homes, although the latter are included in the auto component of the official statistics.

18. This is the mirror image of a relatively slow quarterly rate of adjustment ( $0.13 = 1 - 0.87$ ) to the discrepancy between actual and equilibrium stocks of nonauto durables. See the discussion of this point in Hymans and Shapiro, "The DHL-III Quarterly Econometric Model."

19. In the simulations, *predicted* values of autos and parts are used.

### The Empirical Findings

In discussing the results, I will follow the least common denominator of disaggregation among the models and thus divide the discussion into (1) nondurable goods and services, (2) nonautomobile durables, and (3) automobiles.

#### NONDURABLES AND SERVICES

Despite the differences in the explanatory variables of the statistical relationships and in the degree of disaggregation, the four models yield remarkably similar estimates of the direct impact that the tax surcharge should have had on consumer spending for nondurables and services on the assumption of full effectiveness. They imply that the first-round impact of the surcharge should have curbed consumer expenditures on nondurables and services by between \$1.2 billion and \$1.6 billion at annual rates in the second half of 1968 and by between \$2.5 billion and \$2.6 billion in 1969 (see the impact column in Table 2). The impact estimates fan out a little in 1970, reflecting the varying importance attributed to longer-lagged effects; the largest estimated impact during the first three quarters is \$2.3 billion (annual rate) for DRI and the smallest is \$1.6 billion for Wharton. (All dollar figures throughout are in 1958 prices.)

The lags in the models allow for a considerable buildup period before changes in income have a nearly proportional effect on outlays for nondurables and services. In the first half-year of the surcharge, only about one-fourth of the \$5.5 billion rate of the surcharge is expected to show up as a decline in outlays for nondurables and services. Even during 1969, the expected impact is approximately 35 percent of the \$7.3 billion withdrawal from income accounted for by the surcharge. On the average, however, nondurable goods and services account for about three-fourths of total disposable income. If a change in income were sustained for an indefinite period, the ultimate impact on nondurables and services would be roughly three-fourths in both the Michigan and Wharton models. In the case of OBE, it would be about a half, whereas in the DRI model, it approaches 0.9. Thus, standard econometric models reflect the evidence that sudden changes in income—whether permanent or temporary—have gradually

**Table 2. Impact and Prediction Errors of Selected Models Assuming Full Effect and Zero Effect of Personal Tax Surcharge on Consumption, 1968:3-1970:3**

Billions of 1958 dollars, seasonally adjusted annual rates

Year and quarter	Data Resources, Inc.			Michigan			Office of Business Economics			Wharton		
	Impact error	Full effect error	Zero-effect error	Impact error	Full effect error	Zero-effect error	Impact error	Full effect error	Zero-effect error	Impact error	Full effect error	Zero-effect error
1968:3	1.0	0.8	-0.2	1.5	-0.1	-1.6	1.2	-2.0	-3.3	1.1	-1.2	-2.3
4	1.5	-0.9	-2.4	1.8	-1.8	-3.6	1.6	-3.9	-5.5	1.7	-3.2	-4.8
1969:1	2.4	0.2	-2.2	2.8	-0.5	-3.3	2.5	-3.0	-5.5	2.5	-2.2	-4.8
2	2.8	1.0	-1.8	3.0	-0.3	-3.3	2.9	-2.6	-5.4	2.9	-1.5	-4.4
3	2.6	0.8	-1.8	2.3	-1.4	-3.7	2.5	-3.3	-5.8	2.5	-1.6	-4.0
4	2.8	0.6	-2.2	2.4	-1.8	-4.3	2.7	-3.6	-6.3	2.3	-1.6	-4.0
1970:1	2.6	2.2	-0.3	2.0	-0.6	-2.6	2.4	-2.3	-4.6	1.9	-0.1	-2.0
2	2.6	0.6	-2.0	2.0	-3.4	-5.4	2.3	-5.2	-7.5	1.8	-2.7	-4.5
3	1.8	0.5	-1.3	1.0	-3.3	-4.4	1.4	-5.1	-6.5	1.0	-3.8	-4.8
					<i>Nondurable goods and services</i>							
1968:3	0.4	0.6	0.2	0.6	0.5	-0.2	0.6	1.4	0.8	0.5	0.4	-0.1
4	0.2	-0.3	-0.5	0.7	0.2	-0.5	0.7	1.0	0.3	0.6	-0.2	-0.8
1969:1	0.7	0.4	-0.4	1.1	0.9	-0.2	1.0	1.9	0.9	0.9	0.7	-0.2
2	0.6	0.4	-0.2	1.1	1.5	0.4	1.0	3.1	2.1	0.9	1.2	0.3
3	0.5	-1.4	-1.9	0.7	-1.2	-1.9	0.6	1.3	0.7	0.6	-0.8	-1.4
4	0.7	-1.1	-1.8	0.7	-1.3	-2.0	0.6	2.0	1.4	0.6	-0.9	-1.5
1970:1	0.3	-0.1	-0.5	0.5	-0.5	-1.0	0.4	3.4	3.0	0.4	0.4	*
2	0.5	0.4	*	0.8	-1.3	-2.1	0.4	4.0	3.6	0.4	0.3	-0.1
3	*	-0.3	-0.3	0.4	-1.9	-2.2	*	3.7	3.7	*	-0.6	-0.6
					<i>Nonautomobile durables</i>							
1968:3	0.4	2.2	1.9	0.0	3.7	3.7	0.4	5.1	4.7	0.4	2.5	2.1
					<i>Automobiles</i>							

4	1.4	1.9	0.5	1.0	3.5	2.6	0.4	3.4	3.0	0.5	1.0	0.5
1969:1	0.8	2.3	1.5	0.7	4.2	3.5	0.6	4.0	3.4	0.7	1.4	0.7
2	1.3	2.7	1.4	1.0	4.4	3.4	0.6	4.2	3.6	0.7	1.5	0.8
3	0.5	1.3	0.9	0.7	4.1	3.5	0.4	3.1	2.7	0.4	0.7	0.3
4	-0.1	0.6	0.8	*	3.6	3.7	0.4	3.4	3.1	0.4	0.7	0.3
1970:1	0.3	-1.5	-1.8	0.2	0.2	*	0.2	0.2	*	0.3	-2.8	-2.8
2	-0.1	-0.3	-0.2	-0.1	1.9	2.0	0.2	2.3	2.1	0.3	-0.2	-0.5
3	*	-0.4	-0.4	0.1	0.8	0.7	*	1.6	1.6	*	-0.9	-0.9
<i>Total consumption except for autos</i>												
1968:3	1.5	1.5	*	2.1	0.4	-1.8	1.8	-0.6	-2.5	1.6	-0.8	-2.4
4	1.7	-1.2	-2.9	2.6	-1.6	-4.2	2.3	-2.9	-5.2	2.3	-3.4	-5.7
1969:1	3.1	0.6	-2.5	3.9	0.4	-3.5	3.5	-1.0	-4.6	3.4	-1.6	-5.0
2	3.4	1.5	-2.0	4.1	1.2	-2.9	3.9	0.5	-3.3	3.8	-0.2	-4.1
3	3.1	-0.6	-3.7	3.0	-2.6	-5.6	3.1	-2.0	-5.1	3.0	-2.4	-5.4
4	3.5	-0.5	-4.0	3.2	-3.1	-6.3	3.3	-1.6	-4.9	2.9	-2.6	-5.5
1970:1	2.9	2.1	-0.8	2.5	-1.2	-3.6	2.8	1.2	-1.6	2.3	0.3	-2.0
2	3.0	1.0	-2.0	2.9	-4.6	-7.5	2.8	-1.2	-3.9	2.2	-2.4	-4.6
3	1.8	0.2	-1.7	1.4	-5.2	-6.6	1.4	-1.4	-2.8	1.0	-4.4	-5.4

*Total consumption*

1968:3	1.9	3.7	1.9	2.1	4.1	1.9	2.2	4.4	2.2	2.0	1.7	-0.3
4	3.1	0.7	-2.4	3.5	1.9	-1.6	2.7	0.5	-2.2	2.7	-2.4	-5.1
1969:1	3.9	2.8	-1.1	4.6	4.6	*	4.1	3.0	-1.2	4.1	-0.2	-4.3
2	4.7	4.1	-0.5	5.1	5.6	0.4	4.5	4.7	0.3	4.5	1.2	-3.3
3	3.5	0.7	-2.9	3.7	1.6	-2.1	3.5	1.1	-2.4	3.4	-1.7	-5.1
4	3.3	0.1	-3.2	3.1	0.6	-2.6	3.7	1.8	-1.8	3.4	-1.8	-5.2
1970:1	3.2	0.6	-2.6	2.7	-0.9	-3.6	3.0	1.4	-1.7	2.6	-2.2	-4.8
2	2.9	0.7	-2.2	2.8	-4.7	-5.5	3.0	1.1	-1.9	2.5	-2.6	-5.1
3	1.9	-0.2	-2.0	1.5	-4.3	-5.9	1.4	0.2	-1.2	1.0	-5.2	-6.3

Sources: Derived from data provided by Data Resources, Inc., the University of Michigan's Research Seminar in Quantitative Economics, the Office of Business Economics of the U.S. Department of Commerce, and the University of Pennsylvania's Economic Research Unit. Figures are rounded and may not add to totals.  
\* Less than \$0.05 billion in absolute value.

increasing effects on consumption. The estimated impacts for 1970, which linger on after the surcharge expires, also reflect the diffusion of the effects over time.

The errors on nondurable goods and services in the experimental predictions are clearcut (see Table 2). For every single quarter in all four models, the zero-effect view of the surcharge overpredicts expenditure on these items. However, in three of the models (DRI is the exception), the full-effect view also overpredicts the level of expenditures on nondurables and services in every single quarter, although these errors are substantially smaller in absolute size. In the case of DRI, the full-effect view tends to underpredict a little—on the average, \$0.6 billion in contrast with the \$1.6 billion average overprediction of the zero-effect view. By a more refined statistic described in the appendix and shown in Table 3, the DRI verdict is that the tax surcharge was 69 percent as effective as a standard change in disposable income in curbing consumer expenditures on nondurables and services. By the other three verdicts, however, the surcharge was (at least!) 100 percent as effective as a standard change in income on these components of consumption.

Three of the four models produce errors on nondurables and services during the surcharge period that must be regarded as biased: They clearly do not lie symmetrically around zero. But the errors based on the full-effect view are not large in magnitude. OBE is the only model that ever makes an

**Table 3. Degree of Effectiveness of the 1968 Tax Surcharge on Consumption, 1968:3–1970:3, According to Selected Models**

<i>Type of expenditure</i>	<i>Data Resources, Inc.</i>	<i>Michigan</i>	<i>Office of Business Economics</i>	<i>Wharton</i>
Nondurable goods and services	0.69	1.58	2.46	1.86
Nonauto durables	1.29	1.11	-2.05	0.69
Autos	-0.86	-3.62	-6.85	-0.69
Total nonauto consumption				
Without constraint	0.82	1.48	1.31	1.62
With 0,1 constraint <sup>a</sup>	0.74	1.00	0.78	0.93
Total consumption				
Without constraint	0.53	0.47	0.34	1.33
With 0,1 constraint <sup>a</sup>	0.63	0.88	0.70	0.80
With auto adjustment <sup>a</sup>	0.59	0.88	0.70	0.80

Source: Based on methods set forth in text of appendix; effectiveness is computed over the period 1968:3–1970:3 using tabulations set forth in Table 2.

a. For explanation of constraints and adjustments, see text.

**Table 4. Root-Mean-Square Errors of Full-Effect and Zero-Effect Simulations**

Billions of 1958 dollars

<i>Type of expenditure</i>	<i>Data Resources, Inc.</i>		<i>Michigan</i>		<i>Office of Business Economics</i>		<i>Wharton</i>	
	<i>Full effect</i>	<i>Zero effect</i>	<i>Full effect</i>	<i>Zero effect</i>	<i>Full effect</i>	<i>Zero effect</i>	<i>Full effect</i>	<i>Zero effect</i>
Nondurable goods and services	1.0	1.8	1.9	3.7	3.6	5.7	2.2	4.1
Nonauto durables	0.7	0.9	1.1	1.4	2.7	2.2	0.7	0.8
Autos	1.7	1.2	3.3	2.9	3.3	3.0	1.5	1.3
Total consumption except autos	1.2	2.5	2.8	5.0	1.5	4.0	2.4	4.6

Source: Computed for the period 1968:3 through 1970:3 from data in Table 2.

error as large as 1 percent of the total of consumer outlays on nondurables and services. While there is no standard statistical measure of par for such dynamic simulations, the root-mean-square errors for nondurables and services (Table 4) must be regarded as quite respectable.<sup>20</sup> The errors of the zero-effect view are nearly twice as large as those of the full-effect view. Static errors for nondurables and services (available from the author on request) tend to correct the overpredictions as time goes on; but all except DRI still contain a mean overprediction based on the full-effect view and, of course, larger overpredictions on the zero-effect view. Although they blur the distinctions, the static residuals are consistent with the general verdict in favor of the full-effect view for expenditures on nondurables and services during the surcharge period. They also indicate that the results cannot be attributed to any peculiarity of the dynamic simulation nor specifically to the choice of 1967:2 as the last known quarter. (Further evidence on this issue is presented in the appendix.)

20. "Root-mean-square error" is a standard statistical concept of average error that gives extra weight to particularly large deviations. The errors are squared, and the square root of the average squared error is taken. In a dynamic simulation over a considerable period, the root-mean-square error is expected to be larger than the standard error over the sample period, because the dynamic simulation does not get the benefit of correction from actual lagged expenditure. Standard errors during the sample period for nondurables and services are in the neighborhood of \$1.2 billion to \$1.5 billion. The root-mean-square errors during the simulation period for the full-effect view in three of the models are roughly one to one and one-half times that size—surely as good as should be expected.

Still, it is puzzling that three of the four models overpredicted nondurables and services consistently during the surcharge period. There is no evident built-in tendency for the models to “run uphill” on these components. I have some suspicion that the difference between DRI and the other models may be related to the use of current-dollar variables by DRI. As a percentage of disposable income, outlays on nondurables and services for 1968:3–1970:3 were about the same in current dollars as in the period 1965:1–1968:2; but they were about 0.7 percentage point lower in constant dollars. Price variables could in principle take care of this difference, but it is not obvious that they would.

Three of the four consumption models suggest that any puzzle about consumer spending on nondurables and services during the surcharge period is why it was so *low*. The surcharge is one element that was supposed to hold consumption down, and it does not look at all suspect from this point of view. The models do not reveal what else could be accounting for the extra downward tug on consumption, and they necessarily leave open the possibility that the surcharge may have been ineffective but made to look good by other forces that exerted very large downward pulls on consumer demand. In summary, the data provide no reason for questioning the effectiveness of the surcharge on those components of consumption where the basic challenge of the permanent income hypothesis was focused.

#### NONAUTO DURABLES

As recorded in Table 2, the expected impact of the surcharge on nonauto durables expenditure is quite similar among the models, with the 1969 impact ranging from a low of \$0.6 billion for DRI to a high of \$0.9 billion for Michigan. For the first half of 1970, the expected impact ranges from \$0.4 billion to \$0.6 billion. Lags are short for this component; and, on the full-effect view, the surcharge is expected to curb these outlays a little more than proportionately.

The OBE model consistently underpredicts nonauto durables in every quarter for both the full-effect and the zero-effect versions.<sup>21</sup> The other

21. Curiously, the errors for nondurables and services and those for nonauto durables in the OBE model are strongly offsetting. The combined OBE estimates (covering consumption excluding autos) are remarkably accurate on the full-effect view, as Table 4 reveals.

three models hand down a quite different verdict, however; they consistently overpredict on the basis of the zero-effect view (see Table 2). Of the nine quarters of observations, overpredictions occur in eight cases each for DRI and Michigan and seven for Wharton. On the other hand, in these three models, the full-effect view gives results that are essentially unbiased: The signs of the errors are split 5 to 4 consistently, and the mean error over the nine quarters ranges between a negative \$0.3 billion for Michigan to a positive amount of less than \$0.1 billion for Wharton. As shown in Table 4, the root-mean-square errors for the models other than OBE are not far larger than the roughly \$0.5 billion standard errors of the sample period; moreover, they are uniformly smaller for the full-effect than for the zero-effect view, although not by a wide margin. The measured degree of effectiveness (Table 3) in these three models is close to unity, again consistent with the full-effect view.

#### AUTOMOBILES

According to the full-effect view as interpreted by the four models, the surcharge should have had a prompt and marked impact on automobile demand, curbing outlays by between \$0.5 billion and \$0.6 billion in 1969. But there is no evidence whatsoever that it had any impact. Indeed, even the zero-effect view consistently underpredicts automobile outlays during the second half of 1968 and every quarter of 1969. Only in 1970 do any overpredictions appear in any of the four models: DRI and Wharton then overpredict on both bases, while OBE and Michigan continue to underpredict. But by 1970, the expected impact of the surcharge on autos, on the full-effect view, is negligible—even in a few cases, paradoxically stimulative because of auto stock or income-change effects.

As shown in Table 2, the underpredictions of OBE and Michigan are consistently in the range of \$3 billion to \$5 billion for the second half of 1968 and throughout 1969. For the entire surcharge period, Michigan's and OBE's root-mean-square errors (Table 4) are three times the size of the \$1 billion standard error during the sample period. Wharton and DRI have much smaller root-mean-square errors. They also underestimate during the first six quarters of the surcharge period, but by a considerably smaller average: \$1.8 billion for DRI and \$1.3 billion for Wharton.

In neither the Wharton nor the DRI equation does the size of the auto-

mobile stock have a significant effect on current demand for autos; that is probably one reason for their smaller underpredictions. The rapid buildup of the stock in 1964–68 tends to generate low estimates of demand for 1968 and 1969 in the OBE and Michigan models.<sup>22</sup> The stock market variable of DRI and the interest rate structure variable of Wharton may also help; neither Michigan nor OBE has any credit or financial variable in the auto equation. Finally, the inclusion of much of this period in the sample period for Wharton and DRI may have improved their estimates. It is beyond the scope of this paper to explore more fully the possibilities that might have accounted for the consistent underprediction of automobile demand in the second half of 1968 and in 1969. On the basis of the test of this paper, I find absolutely no evidence that automobile demand was held down by the surcharge.

#### SUMMARY OF EMPIRICAL FINDINGS

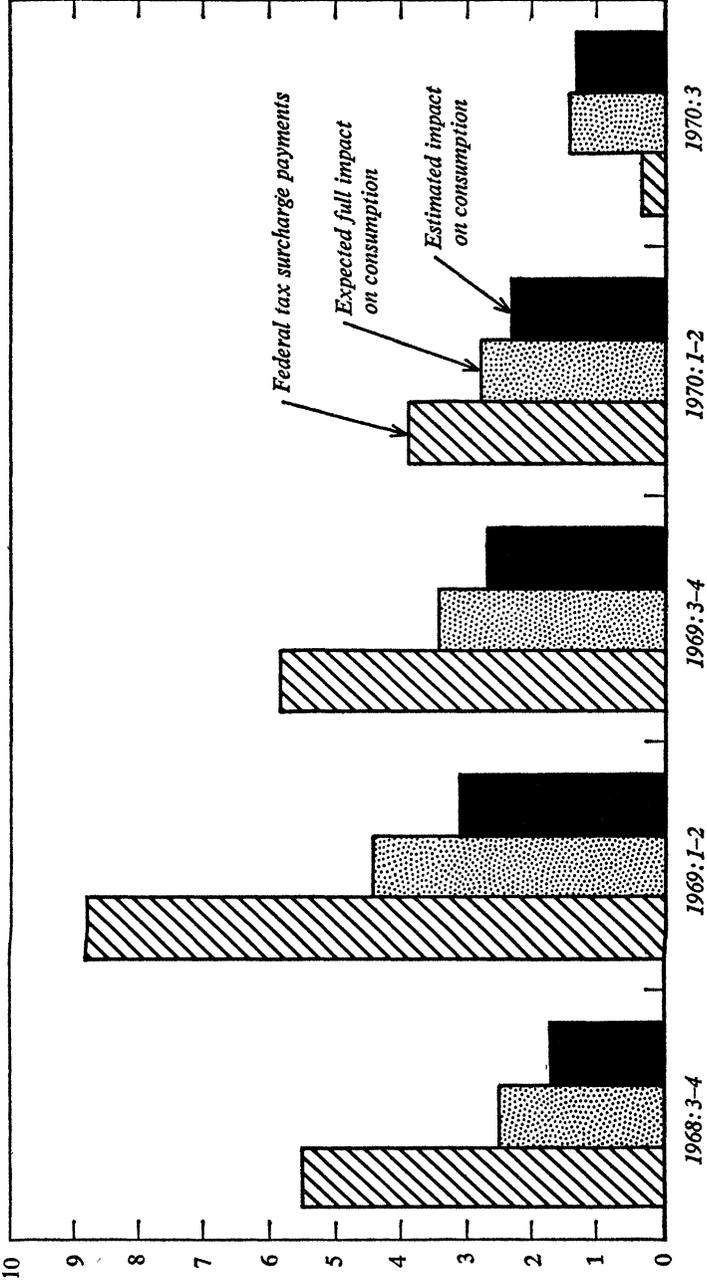
The four models agree remarkably on the amount that the surcharge “should” have curbed consumption in accord with the full-effect view. The direct impact on real consumer outlays was expected to be about \$2.5 billion (annual rate, as always in 1958 prices) in the second half of 1968, \$4 billion in 1969, and between \$2.5 billion and \$3 billion (annual rate) in the first half of 1970. During the entire interval, a little more than half of the surcharge’s drain on disposable income was expected to be reflected in consumer outlays, with a prolonged small effect continuing after the expiration of the surcharge. The expected full-effect impact, averaged for the four models, is shown in Figure 2.

The divergence between the results on automobile outlays and those on other consumption complicates the task of drawing conclusions about the actual effectiveness of the surcharge. On the average, about 14 percent of the expected direct impact was supposed to fall on automobiles, but no evidence was uncovered that automobile demand was, in fact, held down by the surcharge. On the other hand, when consumption of nondurables and services during the surcharge period is interpreted in light of the full-effect view, demand appears to be fairly weak rather than strong. Hence the evidence offers no reason to doubt the effectiveness of the surcharge

22. See the discussion of this point by Saul H. Hymans, “Consumer Durable Spending: Explanation and Prediction,” *Brookings Papers on Economic Activity* (2:1970), pp. 191–92.

**Figure 2. Direct Impact of Tax Surcharge**

*Billions of 1958 dollars, seasonally adjusted annual rate*



Source: Author's estimates; see text.

in that broad area. According to three of the four models, the surcharge was (at least) 100 percent effective on nondurable goods and services in the sense that it had as much impact per dollar as would have been expected from the loss of an equal amount of other income. However, in the fourth model (DRI), the surcharge is estimated as 69 percent effective on nondurable goods and services. The clear verdict on nondurables and services is of particular interest, because the transitory income view of the surcharge points to a low degree of effectiveness for these components.

For nonauto durables, the verdict is essentially the same. Three of the models show a degree of effectiveness ranging between 0.7 and 1.3, while OBE shows no restraint. I find it reasonable to sum up these results by imposing theoretical restrictions on the estimated effectiveness of each component. On analytical grounds, I am unwilling to believe that the surcharge could have had greater than full effectiveness (that is, more impact per dollar than a loss of other income) in curbing consumption excluding automobiles. But neither can I believe that the surcharge actually had a *stimulative* effect on automobile demand. In line with these theoretical beliefs, I take as the summary measure of overall effectiveness a weighted average of the estimated impact on nondurables and services, on nonauto durables, and on autos, but do not allow the estimate for any component to exceed 100 percent or to be less than zero.<sup>23</sup> Then, the estimated effectiveness on nondurables and services is taken as unity for the Michigan, OBE, and Wharton models, and at 0.69 for DRI. The estimated effectiveness for nonauto durables is then zero for OBE, unity for DRI and Michigan, and 0.69 for Wharton. The estimated effectiveness for autos is zero in all four models. The appropriate weight of autos in the average is the expected direct impact on that component as a fraction of the total direct impact on consumption over the period; it is 16 percent for DRI, 12 percent for Michigan, 12 percent for OBE, and 14 percent for Wharton. Allowing for this dilution, and some dilution on nonauto outlays, the effectiveness of the surcharge on total consumption is found to be 63 percent for DRI, 88 percent for Michigan, 70 percent for OBE, and 80 percent for Wharton.

23. Imposing these constraints on each of the three components involves some danger of biasing results. If the surcharge were, in fact, 100 percent effective, purely random errors would result in the estimated effectiveness for some components exceeding 100 percent and others lying below 100 percent. If no figure above 100 percent were accepted, the result would be biased downward. Similarly, if the true degree of effectiveness were zero, ignoring negative values for some components while accepting positive ones would bias the result upward.

If no constraints are placed on the results for the components, then one is obliged to argue that the surcharge had a stimulative effect on automobile demand. Indeed, the result would interpret each dollar of surcharge as equivalent to a *rise* in disposable income of between 69 cents (Wharton) and \$6.85 (OBE) in its impact on auto demand. By the same token, one has to defend Michigan, OBE, and Wharton verdicts that the surcharge was 148, 131, and 162 percent, respectively, effective in curbing nonauto consumption. I find such an interpretation absurd. For example, that approach would have ruled the surcharge 100 percent effective if it had zero indicated effectiveness on nonauto consumption and a coefficient of 800 percent effectiveness on autos. For the same reason that such a result could not be taken seriously, neither can the negative coefficient actually found for autos. Just for the record, nonetheless, the no-constraint approach yields coefficients of overall effectiveness of 34 percent for OBE, 47 percent for Michigan, 53 percent for DRI, and 133 percent for Wharton.

While the extraordinarily strong demand for automobiles cannot reasonably be interpreted as a result of the surcharge, the weakness of demand for other consumption categories might conceivably have been, to some degree, a mirror image of the intensity of automobile demand. Consumers may have held down other expenditures in order to fulfill their desire to spend more on cars. In that event, the strength of automobile demand would have worked along with the surcharge to curb outlays for other components of consumption.

In fact, DRI has investigated the relationship between errors on automobile demand and other components and has found that unexpectedly high expenditures on automobiles are normally associated with especially low outlays on nonauto durables. According to their results, an extra dollar spent on cars above normal rates is associated with 30 cents less spending on nonauto durables and 70 cents less saving. If this adjustment is made during the surcharge period to the estimated demand for nonauto durables, the coefficient of overall effectiveness is reduced from 63 to 59 percent. But the adjustment actually worsens the predictions of nonauto durables within the surcharge period, raising the root-mean-square error from \$0.7 billion to \$0.9 billion.

One implication of the mirror image hypothesis would be that the quarters during the surcharge period when automobile demand was especially strong should have been the quarters in which nonauto consumption demand was especially weak. For Michigan, the reverse was true: Errors in

autos and errors in nonauto consumption were positively related. For Wharton, there was no relationship between the two sets of errors. However, DRI and OBE both displayed a modest negative relationship. The DRI result was consistent with the adjustment made above. The OBE result for the surcharge period suggests that of an extra dollar of auto outlays, 77 cents comes out of saving and 23 cents out of nonauto consumption. If this information is used to generate a new series of OBE predictions for nonauto consumption, assuming that automobile demand is known in advance, the OBE coefficient of effectiveness for nondurables and services remains above 100 percent. The results on the surcharge are thus quite firm, although the behavior of automobile demand remains enshrouded in mystery. The estimated effectiveness based on the auto adjustment line of Table 3 was translated into dollars and averaged for the four models; it is shown in Figure 2, and is close to the expected impact.

In conclusion, the evidence of the surcharge period as interpreted by four econometric models indicates that the surcharge curbed consumption nearly as much as was expected in the models, and that any shortcomings in its effectiveness have no evident connection to the permanent income hypothesis.

### **Perspective on the Results**

These detailed statistical findings stand in marked contrast to some intuitive conclusions that the surcharge was ineffective in curbing consumer outlays.<sup>24</sup> In part, that widespread impression may stem from the unusual strength of automobile demand during much of the surcharge period. Secondly, the reputation may be attributable to the extreme weakness of consumer demand in the spring of 1968, the quarter immediately preceding the initiation of the surcharge. From the second to the third quarter, consumer outlays of all types rose unusually sharply. But the second quarter had been an especially dismal period for soft goods. All the models over-predicted the sum of nondurables and services substantially in that quarter: DRI by \$1.3 billion; Wharton by \$2.4 billion; Michigan by \$2.7 billion, and OBE by \$3.8 billion. During the spring of 1968, the press reported the

24. On the other hand, for two professional views consistent with the findings of this study, see Paul W. McCracken, Statement before the House Ways and Means Committee, May 20, 1969, and Murray L. Weidenbaum, "Fiscal Policy for a Period of Transition," *Federal Reserve Bank of St. Louis Review*, Vol. 52 (November 1970), pp. 14-17.

disappointing performance of retail sales, attributing it in part to the assassination of Martin Luther King and subsequent civil disturbances in several urban areas. Nonetheless, these data were fresh in people's minds during the summer of 1968, and the large rises from those pre-surchARGE benchmarks seemed distressing. In retrospect, it is clear that the abnormality of that period with respect to nonauto consumption was not that the third quarter was unusually strong but that the second quarter was especially weak.

Furthermore, the reputation of the personal tax surcharge may have been sullied by unreasonable expectations rather than inadequate accomplishments. And I am personally obliged to concede that unreasonable expectations may have stemmed in part from a possible overselling of the measure by officials of the Johnson administration who had to convince a reluctant, recalcitrant Congress and skeptical public of the need for fiscal restraint. It may have also stemmed from overly eager journalists who dissected retail sales reports during the summer of 1968 in an effort to reach a verdict on the success of the surcharge. The expected direct impact of the surcharge in that quarter was so small that it could not have been visible in monthly retail sales data.<sup>25</sup>

Another source of intuitive disappointment may be the tendency to focus on the saving rate—the ratio of personal saving to disposable personal income—as the key indicator of the strength of consumer demand. This practice implicitly translates changes in income into expected changes in consumption immediately, unlike the time-consuming process that is embodied in the econometric models. In the models, any change in income has gradually increasing effects on nondurables and services; a year after an income change has begun, only 50 to 60 percent of it is expected to show up in consumption.<sup>26</sup> The models teach the important general lesson that

25. On September 18, 1968, I said: "My statistical estimates would project [the surcharge's] impact on GNP during the third quarter at \$2 billion. . . . There is no evidence yet visible—even in the CEA microscope—that the surcharge has moderated the pace of consumer spending. But there was never any reason to believe that there should be evidence of this nature by now." Arthur M. Okun, "Perspectives on Business in 1969," *Conference Board Record*, Vol. 5 (November 1968), p. 22.

26. The single aggregate consumption equation I employed in my 1965 paper on the tax cut worked somewhat faster than these models; it implied that, four quarters after an income change was initiated, consumption was affected by 80 percent of the change in income. I am not totally convinced that the lag structure of these models is accurate; indeed, too slow a process could be one source of the overpredictions of nondurables and services during the surcharge period.

it takes time for tax changes—whether permanent or temporary—to build up to their complete direct impact on consumption.

By the verdict of the models, however, the personal tax surcharge was no placebo. In curbing consumption directly by most of the expected \$4 billion (1958 prices) in 1969, the surcharge also must have triggered off important multiplier and accelerator effects, and must have had significant induced effects on prices by dampening demand. It can easily be credited with curbing the growth of money GNP by at least 1 percent during its initial year of operation.

Nonetheless, both real and money GNP continued to exceed policy targets until late in 1969. As discussed above, automobile demand displayed surprising strength. Business investment rose sharply, defying all forecasts. Homebuilding meanwhile proved to be resistant to rises in interest rates. Two weak monthly reports on housing starts for May and June 1968 (which subsequently turned out to be erratic) contributed substantially to the misassessment of the true strength of private demand and to the mistakes in monetary policy in the summer of 1968. Later on, when the boom was finally halted by a combination of fiscal and monetary restraint, inflation proved to be remarkably immune to market forces.

All of this history merely documents a misdiagnosis of the economic situation. But it is important to recognize the nature of the errors in economic diagnosis and prescription. According to the evidence cited above, the medicine of the personal tax surcharge did lower the patient's fever. To be sure, the patient was more feverish than the doctors recognized and consequently their antifever prescription was inadequate. But don't blame the medicine; it did most of what it should reasonably have been expected to do. In short, the evidence of the surcharge period provides further confirmation of the general efficacy and continued desirability of flexible changes in personal income tax rates—upward or downward, permanent or temporary.

## APPENDIX\*

THIS APPENDIX CONTAINS TWO PARTS. The first explains the derivation of the coefficient of effectiveness statistic shown in Table 3 and used in the text. The second part discusses the significance of the use of 1967:2 as the last known quarter in the dynamic simulation. A supplementary appendix, available upon request, gives the prediction errors resulting from static simulations of the four econometric models for the surcharge period and lists their equations of the consumption sector.

**Measurement of Degree of Effectiveness**

Suppose the surcharge had some impact per dollar in curbing consumption but only a fraction  $a$  ( $0 < a < 1$ ) as much as a dollar loss of income from other sources. The "proper" income variable is then neither official disposable income ( $Y$ ), which is the choice of the full-effect view, nor the zero-effect concept that adds back the surcharge ( $S$ ) to get  $(Y + S)$ , but something in between:  $Y + (1 - a)S$ . If  $M$ , the impact calculated according to the full-effect view, is linear and homogeneous in current and lagged  $S$ , then the impact attributed to the surcharge in the partial-effect view is  $aM$ . The predicted value of consumption in the partial-effect view will be below that of the zero-effect view by  $aM$ , and the residual ( $W$ ) will exceed the zero-effect residual ( $V$ ) by  $aM$ :

$$(1) \quad W = V + aM.$$

Over the period of observations, the sum of squared errors of predictions based on the partial-effect view is

$$(2) \quad \sum W^2 = \sum V^2 + 2a \sum MV + a^2 \sum M^2.$$

To determine  $a^*$ , that value of  $a$  most consistent with the evidence, one should minimize the sum of squared errors in (2). First differentiate (2) with respect to  $a$ :

$$(3) \quad \frac{d(\sum W^2)}{da} = 2 \sum MV + 2a \sum M^2.$$

\* This appendix was prepared by Richard H. Mullins.

Assuming the existence of an extremum (necessarily a minimum), set (3) equal to zero and solve for  $a^*$ :

$$(4) \quad a^* = \frac{-\sum MV}{\sum M^2}.$$

Figures in the text and in Table 3 referring to the "degree of effectiveness" are values of  $a^*$ .

The OBE, Wharton, and Michigan models are in accord with the assumption that  $M$  is linear and homogeneous in  $S$ ; DRI departs from it in only one trivial respect, through its percentage income-change variable in the auto equation.

The calculated  $a^*$  is not constrained to lie between zero and unity. If the  $V$ s tend to be positive (that is, if even the zero-effect view underestimates consumption), then the calculated degree of effectiveness will be negative:  $a^* < 0$ . On the other hand, if the  $V$ s are negative and tend to be larger in absolute value than  $M$ , so that even the full-effect view generally overestimates consumption,  $a^*$  will turn out to be greater than unity.

### **Choice of the Last Known Quarter**

As discussed in the main body of the paper, the technique of dynamic simulation requires that the investigator be "blindfolded" with respect to information on actual consumption expenditures after some point. From that point on, he is compelled to use his own predicted values of consumption whenever lagged consumption is an explanatory variable in an equation. It was decided in advance to use 1967:2 as the last quarter for which actual consumer spending was to be known. After the simulations were completed, a check was made on the Michigan and Wharton models to see how other choices for the last known quarter would have altered the results.

Whenever lagged consumption is an explanatory variable, the strength of demand (relative to that predicted by the equation) in the last known quarter affects the subsequent track of predicted values. If consumer outlays are especially strong in that quarter, then the predicted value for the next quarter is higher because it includes some positive coefficient multiplied by the expenditure of that last known quarter. If the coefficient on lagged consumption expenditure is far below unity, the effect of strength in the last known quarter will decay rapidly. But if it is as high as 0.9, as

it is in some of the equations for nondurable goods and services, then the residual of the last known quarter exerts a substantial impact on the track of predicted values for a considerable period.

None of the four models used in this study was very far from the mark on nondurables and services in 1967:2, although all four overpredicted: Wharton by \$0.4 billion, DRI by \$0.7 billion, Michigan by \$1.1 billion, and OBE by \$1.7 billion. In the case of nonauto durables, Michigan is the only model in which lagged expenditure enters significantly as an explanatory variable; the Michigan model predicted nonauto durables exactly in 1967:2. In the case of autos, lagged expenditure is included in the Michigan and OBE models in a significant way; both underpredicted this component in 1967:2, Michigan by \$0.9 billion and OBE by \$2.1 billion.

The use of every quarter from 1966:1 to 1968:2 as the last known quarter was explored for Michigan. In the case of automobiles, the choice turns out to have no significance at all: Every one of the ten hypothetical jumping-off points leads to exactly the same errors to the nearest tenth of a billion dollars after 1969:1; at the start of the surcharge period in 1968:3 the biggest difference of any simulation from that employed in this paper is only \$0.3 billion. This result reflects the fact that the coefficient on lagged automobile expenditures is 0.34, small enough to decay rapidly. In the case of nonauto durables, the highest track of predicted values would have been obtained with 1966:3 as the last known quarter. In 1968:3, that track is \$0.4 billion above the one actually employed with the 1967:2 jumping-off point; but the difference diminishes to only \$0.2 billion in 1969:3, and is down to \$0.1 billion by 1970:3. On the other extreme, the lowest track of predicted values would have been obtained with 1967:3 as the jumping-off point, and that track differs from the one actually used by \$0.2 billion in 1968:3 and only \$0.1 billion in 1970:3. None of these alternative choices would have made a qualitative difference in the findings.

In the case of nondurable goods and services, the various tracks of predicted values differ substantially and remain far apart throughout the surcharge period. The choice of 1966:1, 1966:2, 1966:3, and 1967:1 as the last known quarter would have yielded an even higher track of predicted values than did that actually employed and thus would have resulted consistently in even larger overpredictions of actual spending on nondurables and services. The choice of 1968:1 as the last known quarter would have yielded a track essentially indistinguishable from that of 1967:2. On the other hand, the choice as the last known quarter of weaker

quarters of consumer spending on nondurables and services, such as 1966:4, 1967:3, 1967:4, and 1968:2, would have resulted in lower tracks of predicted values than did that actually employed in the dynamic simulation.

The only jumping-off point that would have changed Michigan's basic conclusions on nondurables and services would have been 1968:2. Reflecting the extreme weakness of consumer demand for these items during that quarter, the subsequent track of predicted values for the full-effect view would have been enough lower to produce an average *underprediction* of consumer spending on nondurables and services during the surcharge period. Quite apart from the extreme weakness of consumer demand during that quarter, any economist planning such an experiment in advance would have decided not to use the quarter immediately preceding the surcharge as the last known quarter for dynamic simulation just because it would emphasize whatever errors or shocks occurred in that quarter. The rationale of any dynamic simulation argues for getting a head start several quarters before the surcharge period. The next lowest track of the ten possibilities explored took 1967:4 as the last known quarter; on the full-effect view, it would have shown nearly unbiased results with an average underprediction of \$0.2 billion a quarter during the surcharge period. Every other possible jumping-off point that was explored yielded substantial average overpredictions, as did the one actually employed.

The Wharton model has a coefficient of 0.37 on lagged expenditures for nondurables, and this effect decays rapidly; no jumping-off point between 1966:1 and 1968:2 produces differences amounting to more than \$0.3 billion in any observed quarter after 1968:3. In the case of services, which have a coefficient on lagged expenditures of 0.98 in the Wharton model, a few of the tracks are considerably different from the one actually employed with the 1967:2 jumping-off point. The highest track of predicted values is that employing 1966:1 as the last known quarter: At the start of the surcharge period in 1968:3, it is \$0.6 billion above the track actually employed, and at the end of the surcharge period in 1970:3 it is still \$0.5 billion higher than the track used above. The lowest track of predicted values takes 1968:2 as the last known quarter; it runs consistently about \$1.2 billion below the track that was actually employed. But even that lowest track of the ten yields an average overprediction of service outlays on the full-effect view, overpredicting in six of the nine quarters of the surcharge period.

## *Comments and Discussion*

**Saul Hymans:** The obvious hallmarks of Okun's paper are analytical care and clarity of presentation. That makes my task extremely difficult. In fact, it leaves me with very few options if I want to speak longer than it takes to congratulate the author. One option is to deliver a paper of my own, thereby saying to the original author, "You wrote a fine paper, except that it is the wrong one." I will not choose that option, because I believe the right paper was written. Another option is to allow myself a bit of careless abandon in contrast to the analytical care exercised in the paper. Are there then some wild, tentative inferences that might be worth thinking about in our studies of either consumer behavior or fiscal policy?

Let me postulate that it is mid-1968, and consumers are becoming convinced that inflation is serious and growing worse. At the very least, wealth portfolios will have to be adjusted, shifting away from financial assets and toward real goods. An obvious way to do this at the margin is to buy more automobiles at the cost of lower financial savings. Indeed, it might even be necessary to accumulate more wealth by forgoing some expenditures on nondurable goods and services. Now the surcharge is introduced into this environment, and threatens a diminution in the accumulation of wealth precisely at a time when it appears that the composition of wealth is wrong and its rate of growth inadequate for future needs. To cut back on auto expenditures is precisely the wrong response when the need is for more real wealth. In order to prevent any decline in the growth of total wealth, the increase in auto spending is financed by a relative reduction in expenditures on nondurables rather than a relative drop in financial savings.

Suppose this story were true. How does it then appear to the econometric eye trying to track the surcharge? It makes the surcharge look more than

100 percent effective with respect to nondurables and services and perverse with respect to automobile expenditures. Admittedly, this is a story cooked up to explain the observed data. But it is not that farfetched, at least not at the margin. The appropriate test of this tale is not in the quarter-to-quarter movements of automobiles and nondurables, but rather in the nine-quarter trend over the period.

Consumer spending on durable goods is not just another category of consumption. It is a category of wealth accumulation, of investment. It is under the control of those individuals who feel the quick and direct effect of changes in tax policy. We are never quite sure what *business* tax rates are going to do to *business* fixed investment in any short interval of time. And the same is true of the impact of changes in *personal* tax rates on *personal* fixed investment. This uncertainty makes the conduct of fiscal policy much more complicated. It is not possible simply to apply multipliers to estimates of the effect of tax changes in order to derive the economic impact.

**Lester Taylor:** Arthur Okun has done an important and relevant piece of work. In my opinion, the mechanics of the experiment gave the permanent income hypothesis a fair test, or at least as fair a test as is possible. I find convincing his conclusion that, whatever the reasons why consumer spending was not curtailed following the surcharge as fully as expected, they were not such as to confirm the permanent income hypothesis. If consumers spent according to the permanent income hypothesis, and viewed the surcharge as only a transitory reduction in income, then the bulk of its effect should have been seen in durable goods and in the rate of saving. Its observed effects on nondurable goods and services should have been negligible. However, as the results show, neither of these two predictions materialized. On the average, expenditures for nondurable goods and services were pretty much in line with the predictions based on a full effect from the surtax, while expenditures for automobiles and parts actually boomed.

Let me now turn to a few smaller matters. Any consumption function worth its salt these days is dynamic. Typically, this takes the form of including expenditures of the preceding period as a predictive variable. The presence of a lag term tends to impart spurious accuracy to the predicted values because of the positive autocorrelation that is inherent in economic time series. Because of this, Okun properly used dynamic rather than static simulation. However, I would take issue with his interpretation of the role played by the lagged expenditure variable. He states that the

proposition that expenditures in the current quarter for gasoline depend in part on previous expenditures for gasoline does not imply a causal relationship; it implies rather that lagged expenditures are designed to reflect habit formation and the persistence of consumer tastes through time. I think that whether the lagged expenditure term has causal significance depends upon the particular rationale of the model and, incidentally, upon one's view of causality. What I have in mind, continuing the example of gasoline, is that current expenditures would be related to income through a state variable reflecting the force of habit; that state variable is not observable and its place is taken by the expenditures of the last period. On this view these have a causal interpretation. I would also modify his statement that lagged expenditures necessarily reflect habit formation. They can also have other interpretations, which do not suggest short-run inertia in spending. For some commodities, especially durable goods, the existence of inventories is represented by lagged expenditures and results in stock adjustment spending responses as incomes change. Whether or not lagged expenditures stand in for a state variable representing habit or inventories, the result will be a positive coefficient on last period's expenditures. But the coefficient for goods characterized by habit formation and short-run inertia will typically be 0.8 or higher, while the coefficient for goods characterized by inventories and stock adjustments will be on the order of 0.5 or even lower.

**Robert Eisner:** Many of us questioned the surcharge as a device to fight inflation on a number of grounds, some political-economic, some frankly political. The political-economic reasons related to a whole complex of forces. We felt that the main forces pushing toward inflation were the government expenditures for the war. It did not seem possible to divorce tax policy from expenditure policy. A refusal to go ahead with the surcharge might actually have led to cuts in government expenditures or at least restraints on their increase. On the economic side, in the passage Okun quoted from, I argued that "the basic economic error" applied to business investment as well as to consumption. Okun essentially concedes my point on the trivial impact of the temporary corporate surcharge on investment. And he ignores the extension of excise taxes. He chooses to review only one piece of the fiscal program, the personal surcharge.

Let us look at what happened. The GNP deflator rose at an accelerating pace from 1965 through 1970. Surcharge or not, the end of inflation did

not come. Business fixed investment rose from \$81.6 billion in 1966 to \$99.3 billion in 1969. Consumption, which had been rising at a rate of 10.4 percent in the first half of 1968, rose by 8.9 percent in the last half, 7.5 percent in the first half of 1969, and 7.1 percent in the last half of 1969. The personal saving rate, which offers a guide to how consumers are spending, was 7.6 percent of disposable income in the second quarter of 1968, just before the imposition of the surcharge, and dropped to 6.1 percent in the third quarter. In the six quarters before the surcharge, it averaged 7.4 percent; it fell to an average of 6.1 percent the six quarters after the surcharge was imposed. Looked at in this crude fashion, the main effect of the surcharge was simply a reduction in the saving rate.

Now let me focus more narrowly on Okun's paper. He has pointed to the reduction in the real rate of growth of output after the overall stabilization policy turned more restrictive; but one might wonder how much longer the real rate of growth could have been expected to continue at the 6½ percent rate that prevailed in the first half of 1968. Given the ceiling on supply, sooner or later we had to come down to lower rates of growth. In addition, the impact of inflation was uneven, which, one might conjecture, might have had some inhibiting influence on real consumption.

Okun himself points out several flaws in the results of the model. But he is encouraged by the relatively low root-mean-square errors for non-durable goods and services in the prediction period. Nonetheless, the root-mean-square errors are about the same magnitude as the estimated surcharge impacts themselves. The surcharge itself was relatively small in size and the predicted errors about the same size. It is also troublesome that more than half of the unconstrained values for  $a^*$ , the variable developed to measure the degree of effectiveness of the surcharge, fall outside the reasonable a priori bounds of zero to one. I do not think that it is correct to constrain the individual  $a^*$  to these bounds before forming an overall measure of surtax effectiveness. The values given without constraint are a better measure of the effectiveness of the surcharge than the ones with the constraint, and three of these four estimates range from 34 to 53 percent. Okun further concedes that, if he had used 1968:2 as the starting point for the dynamic simulation, his results would have been even closer to the zero-effect view.

One thing that has been overlooked is that the models, in effect, build in the permanent income hypothesis by making current consumption a function not only of current disposable income, but of lagged values of

disposable income or consumption as well. The variations in income over time are in considerable part transitory and that is why the models give a small response to a one-quarter or even a one-year change in income. When consumers experience a change in income and do not perceive it as permanent, they react slowly at first. If the rise in income persists, the consumer forms a firmer notion that the change is permanent and his assets rise and he spends more. During the first four quarters of the surtax I calculate that the models were predicting only 38 cents of reduction in consumption for each dollar of surtax. This may be compared with Okun's suggestion that Friedman expected a 33 cent response on the basis of his permanent income hypothesis. The fact that Okun finds, even with his constraint, that the surcharge is only 59 to 88 percent as effective as the models predicted indicates that the surcharge was viewed as even more transitory than other elements of income.

I might add parenthetically that while I quite prefer the permanent income approach to the "ratchet argument," Okun hardly does justice to the latter. In this context, the ratchet would apply not merely, as Okun suggests, to the "fraction of total income belonging to the group of families that was pushed from no change (or a minuscule rise) in income into the minus zone as the result of the surcharge." It would rather apply as well to all those whose incomes *without the surcharge* were less than their previous peaks. Given both the variability of money incomes and the fact that the real incomes of many households were declining because of inflation, contrary to Okun's conclusion, it is easy to believe that *this* group accounted for "a significant fraction of total income," which would, according to the ratchet argument, be subject to a quite low marginal propensity to consume.

**Arthur Okun:** The expectation that the surcharge might have a dollar-for-dollar impact, a super-full effect overnight, was a mistake in some people's expectations of what it could do. As I said in the paper, I wonder whether the surcharge may have been oversold and people expected too much from it. But that is not the issue here. The question is whether the surcharge, because it was temporary, was less effective in restraining consumption, particularly on nondurable goods and services, than a permanent tax change would have been. What these models tell us is that consumption lags behind income. According to these models, it takes nearly a year before a dollar change in income has a half-dollar direct effect on con-

sumption. The effect is gradual regardless of the kind of income change. The results indicate that this tax increase worked no differently from what the models would have expected on the basis of a permanent tax increase, or an income change arising from any other cause. And it is critical that the permanent income hypothesis would stress nondurables and services as the key components for a test.

I should like to respond to a couple of technical points. On the question of whether to constrain  $a^*$  when adding up a measure of effectiveness, doing otherwise seems absurd to me. Adding up without constraints says that the surcharge stimulated automobile demand and reduced nonauto demand with more than full effect. Applying the zero-to-one theoretical constraint is simply a systematic and symmetrical way of ruling out both these unrealistic interpretations. It should not be bothersome that prediction errors in the consumption function often lead to point estimates of  $a^*$  outside that range. In comparing the errors to the estimated impact, one has to keep in mind the size of the surcharge and of the typical errors of consumption equations. If the surcharge had been larger, or if our econometric models were more accurate, there would have been more discrimination. The fact that the surcharge was not large relative to the normal errors that consumption equations make is good reason to expect a considerable standard error on the  $a^*$  calculation. The estimates of 1.29 and even 1.58 for  $a^*$  are no surprise in this context. But negative numbers of the size of 6.85 and 3.62 for the auto sector are troublesome; they reflect the big errors in predicting autos and cannot be taken seriously. In any case, one cannot complain both about the substantial errors inherent in consumption forecasts and about the fact that  $a^*$  estimates frequently lie outside the theoretical limits of zero and one. These are not two separate problems.

### **General Discussion**

Robert Eisner extended his earlier comments, stressing that it was necessary to consider the nature of any change in income and its effect on expected future income and assets before judging what consumption effect to expect. He contended that a permanent change in taxes might well have a larger and prompter impact than the models would imply. R. J. Gordon noted that the prediction errors of all consumption sectors combined in the surcharge period came out favoring the no-effect view in the first four

quarters and the full-effect view subsequently. He argued that monetary policy was disregarded in the models and that easy money in the first quarters following the surcharge could have explained this error pattern. Perry and Okun agreed that the monetary influences may have been important. These factors provide the basis for an interesting conjecture about what caused auto demand to be so strong in the early quarters. Perry pointed out, however, that there was no special error pattern to be explained for nondurables and services. Okun noted that the Wharton model included the money stock in its consumption equations and still gave answers very similar to those from the other models.

William Fellner and others expressed doubts about the reliability of the results because of the persistent errors in the predictions, particularly in the auto equations. Fellner wondered whether the superiority of the full-effect view over the zero-effect view was sufficiently great to be convincing, even for nondurables and services. Okun agreed that sophisticated reliability tests would be useful, but they were unobtainable for this exercise. He did feel the errors for nondurables and services on the full-effect view were small enough and sufficiently smaller than those of the zero-effect view to be intuitively persuasive. The full-effect proponent wins the bet by a sizable margin.

William Brainard felt uncomfortable that the exercise allowed the surcharge variable to have the first crack at explaining the errors. Ideally, other variables should have been given an equal opportunity to explain them, including a reestimation of coefficients on the original variables to allow for a possible gradual shift in structure.

R. J. Gordon and William Branson believed there might be offsets between spending on autos and on other consumption that would call for more emphasis on the behavior of aggregate consumption. Okun pointed out that such offsetting effects as the models could identify were allowed for in his computations; but the models themselves found that most of the unusual behavior in auto purchases during the estimation period was offset by changes in saving rather than changes in other consumption spending. If much stronger offsetting effects were suspected in the projection period of the surcharge, they should have shown up in the quarterly error patterns; but they did not. Gordon, however, felt that the offsets need not be simultaneous in the way Okun suggested; consumers might hold down their nonauto consumption either before or after buying a car, in order to help finance the downpayment.