A Study to Evaluate the Effect of Reduced Greenhouse Gas Emissions on Vehicle Miles Traveled

Final Report
ARB Contract Number 02-336

Principal Investigators:

Kenneth A. Small, Ph.D.

Department of Economics

University of California, Irvine

Kurt Van Dender, Ph.D.

Department of Economics

University of California, Irvine

March 2005

DISCLAIMER

The statements and conclusions in this report are those of the researchers and not necessarily those of the California Air Resources Board, the California Energy Commission, or the University of California. The mention of commercial products, their source, or their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

List of Figures

LIST OF TABLES

ABSTRACT

This report presents measures of the size of the rebound effect, by which improvements in fuel efficiency of vehicles may cause vehicle travel to increase. We use aggregate cross-sectional time series data for 1966 to 2001 on all 50 U.S. states and the District of Columbia. Our model contains a measure of the historical effects of the federal Corporate Average Fuel Economy (CAFE) standards, which helps stabilize results compared to previous literature. Also, our time series is longer than previous studies, enabling us to better discern the difference between short- and long-run effects. Our best estimate of the rebound effect for the US as a whole, over the period 1966-2001, is 5.3% for the short run and 26% for the long run. We also find that the rebound effect declines with income. Using the 1997-2001 average value of income for California, the short- and long-run rebound effects are estimated at 2.2% and 11.3%, respectively. Our methodology permits projections to future years, including dynamic projections accounting for changes in income occurring at the same time as owners are adjusting from the short to the long run. These results enable researchers to predict how proposed standards for greenhouse gas emissions in California may affect the amount of vehicle ownership and travel.

EXECUTIVE SUMMARY

Background

Under A.B. 1493, the California Air Resources Board is required to regulate greenhouse gas emissions from motor vehicles. Because such emissions regulations result in greater fuel efficiency, the per-mile cost of operating a motor vehicle is reduced. One of the technical issues the Board faces is how much this reduced cost might increase the amount of driving undertaken by California vehicle owners. Although there is past research on this, it is not specific to California and has an uncertain validity when applied to the future time period in which regulations would take effect.

Methods

This report attempts to answer the question of how big the rebound effect is likely to be in California in the near future. It does so by estimating an econometric model, based on 36 years of data in 51 states plus the District of Columbia, explaining the amount of travel by passenger vehicles as a function of cost and other factors. It specifically considers factors that differ between California and other states, such as income and degree of urbanization.

The standard definition of the rebound effect relates travel, usually measured as vehicle miles traveled (VMT), to fuel price and fuel efficiency. Our analysis takes into account that fuel efficiency is chosen by drivers, by postulating a three-equation structural model describing how consumers choose the number of vehicles to own, the amount of driving, and average fuel efficiency. In our theoretical section, we use the structural model to distinguish three channels through which fuel-efficiency choices may be affected: regulatory changes, fuel price changes, and vehicle price changes.

A literature review produces a number of guidelines for the empirical work. First, a reliable estimate of the rebound effect from historical data requires careful accounting for the federal Corporate Average Fuel Economy (CAFE) standards, because they affect fuel-efficiency choices. Second, time-series estimates of the long-run rebound effect are very sensitive to the econometric specification. The most compelling evidence from the literature we reviewed is from combined time-series and cross-sectional studies, which suggest that the long run rebound effect in the US is around 23%.

In combination with the literature review, our review of available data leads us to opt for an econometric estimation of the rebound effect using aggregate cross-sectional time series data. We are able to compile such a data set for 1966 to 2001 on a cross-section of U.S. states and the District of Columbia. Such data provide direct evidence of the aggregate effects that are of policy interest, they enable us to identify California-specific effects, and they allow us to better distinguish among various dynamic postulates in order to estimate more precisely the difference between long run and short run rebound effects.

We formulate a measure of the effects of the CAFE standards as the nation-wide difference between the standards and desired fuel efficiency. Desired fuel efficiency is predicted using an

estimated function based on pre-CAFE data. This CAFE variable seems to help stabilize results, which have shown considerable variation in the literature.

Our specification allows for time-varying elasticities and rebound effects, to reflect the possibility that the rebound effect declines over time. Such a decline has been hypothesized because the share of fuel costs in total per mile costs declines as vehicles become more fuel-efficient and as the opportunity cost of time spent traveling increases, due to rising incomes and more congestion. We capture these effects by allowing the rebound effect to vary with income and degree of urbanization.

One disadvantage of our data set is that fuel efficiency does not vary much over time and space, making it impossible to determine its effect separately from that of fuel prices, which vary a great deal. Therefore our study, like most previous ones, relies on the theoretical assumption that people react the same way to a change in fuel cost whether it arises from a change in price or a change in efficiency. This is a reasonable assumption because fuel costs, unlike costs of operating various electrical appliances, are highly visible as the motorist must refuel regularly and the expense of doing so is clearly displayed. However it is an assumption we are unable to test satisfactorily using these data.

Results

Our best estimate of the rebound effect for the US as a whole, over the period 1966-2001, is 5.3% for the short run and 26% for the long run — the latter value being remarkably close to the consensus from the literature noted above. Furthermore, the rebound effect depends on income in the expected manner, a result not previously measured in the literature. At the value of income equal to the 1997-2001 average in California, the rebound effect is considerably smaller: 2.2% in the short run 11.3% in the long run. Additional estimation results, like the long-run overall price-elasticity of fuel demand (-0.46) and the proportion of it that is caused by mileage changes (55 percent) are very much in line with the literature.

Projections to future years are inherently uncertain because California per capita income is expected to rise to values not seen in our sample. Nevertheless our best effort at making such a projection yields a long-run rebound effect that declines to about 7.8% in 2009 and 4.4% in 2020. We also project a "dynamic" rebound effect, showing as an example that a permanent change starting in 2009 produces an effect on VMT that starts small, grows gradually as behavior adjusts toward a long-run response, but then declines as rising incomes reduce the long-run rebound effect; the maximum rebound effect in this calculation is 5.9%.

Conclusion

Our results show a long-run rebound effect that is quite similar to that of other studies when taken as an average over the US for 1966-2001. Its dependence on income, however, implies that it is considerably smaller for California when projected to the years in which greenhouse gas regulations will take effect. Furthermore, our empirical specification permits ARB staff to make year by year projections for any regulatory scenario they might wish to consider, a process that has already been utilized in formulating background material for the Board's consideration of regulations implementing A.B. 1493.

1.   Introduction

Our research objective is to produce estimates of the rebound effect that are useful for analyzing the impact of policies related to reductions of greenhouse gas emissions from private vehicle transport in California.

The rebound effect (or 'take back effect') is usually defined with respect to some form of energy consumption, such as use of fuel or electricity. It refers to a situation where a regulatory or technological change causes an improvement in the energy efficiency of some equipment (e.g. an air conditioner) that is an input to the production of a desired service (air cooling). This improvement has the side effect of making the service itself cheaper. Assuming a downward-sloping demand curve for the service, the quantity of service demanded then increases (people keep their houses cooler). This in turn causes a rise in the derived demand for the energy input (electricity), thereby offsetting some of the direct effect on energy consumption of the original change. The rebound effect can be defined as the difference between the actual quantity of energy demanded after the change and the hypothetical quantity that would be demanded if the usage of the related service were constant.

For motor vehicles, the energy input is fuel and the associated service is travel, typically measured as vehicle-miles traveled (VMT). In this case the rebound effect arises because when vehicles are made more fuel-efficient, it costs less to drive a given amount so VMT goes up. That in turn causes more fuel to be used than would be the case if VMT were constant; the difference is the rebound effect.

Obtaining reliable measures of the rebound effect is potentially important for choosing policy instruments to reduce emissions of greenhouse gases from motor vehicles. This is because carbon dioxide (CO_2) accounts for a large fraction of the total greenhouse gas emissions from motor vehicles, and there is a close link between CO_2 emissions and fuel consumption. Thus, California regulations to reduce greenhouse gas emissions will likely, as a side effect, result in increases in vehicular fuel efficiency, making it important to better understand possible rebound effects.

Most of the literature on the rebound effect is stated in terms of fuel efficiency. Proposed California regulations will address greenhouse gas emissions rather than fuel efficiency. We have formulated our theory to address any regulatory action that includes a rather direct effect on fuel efficiency, whether or not that is its main intent. Since this is very likely to be a primary aspect of the final regulations, our theory and empirical findings are appropriate to analyze those effects of greenhouse gas regulations that directly affect manufacturers' and consumers' choice of fuel efficiency. For these reasons, we refer to "fuel efficiency" in most of our discussion, bearing in mind that it is one component of the full analysis of greenhouse gas regulations.

If the rebound effect is large, the effectiveness of regulatory actions that improve fuel efficiency is diminished unless they are accompanied by other measures to increase the price of VMT, such as an increased fuel tax. By contrast, if the rebound effect is small, fuel efficiency

standards are relatively effective while the tax increase required to achieve a given objective may be so high as to be politically infeasible (Greening and Greene, 1997).[1]

In addition, fuel efficiency standards or other greenhouse gas regulations are likely to affect the price of new vehicles. This in turn sets up a number of market reactions, including a tendency to reduce the size of the vehicle stock and an increased average age of the stock. We account for the first of these effects through an equation predicting vehicle stock. We do not account for the average age of the fleet or the effects that may have on driving and on emissions; we believe that analysis requires a more micro-based model that explains individual household vehicle purchase decisions.

2. Literature

In this section, we consider in some depth a representative sampling of literature analyzing the rebound effect for fuel use in motor vehicles. Each study is chosen to represent a particular empirical approach. For other literature surveys on the rebound effect and related issues, see Greening et al (2000) and Graham and Glaister (2002).

2.1 Defining the rebound effect

2.1.1 Standard definition

The rebound effect for motor vehicles is typically analyzed in terms of an exogenous change in fuel efficiency, E, measured for example in miles per gallon (e.g. USDOE, 1996). Fuel consumption F (in gallons per year) and travel M (vehicle-miles traveled per year) are related through the identity $F=M/E$. The demand for fuel consumption is therefore naturally viewed as a derived demand, arising from the demand for vehicle-miles traveled (VMT). The latter depends (among other things) on the variable cost per mile of driving, which includes the per-mile fuel cost, $P_M \equiv P_F/E$, where P_F is the price of fuel. Therefore the demand for fuel can be written as:

$$F\left(P_F, E\right) = \frac{\hat{M}\left(P_M\right)}{E} = \frac{\hat{M}\left(P_F/E\right)}{E}. \tag{1}$$

We use the notation $\hat{M}(\cdot)$ for this function as a reminder that it is a very particular form of a demand function: namely, one that depends on fuel price and fuel efficiency only through their ratio, and that does not depend on other things that might be influenced by fuel price and fuel efficiency. This function is derived more precisely in Section 3.1

---

[1] Moreover, the textbook case for defending taxes over standards is less strong when account is taken of some real-world deviations from textbook assumptions such as imperfect competition in the passenger car market and decision-making costs for consumers. See CBO (2002) for a comparative analysis of policy instruments to reduce gasoline consumption.

The simplest version of the rebound effect follows immediately from (1). The elasticity of fuel demand with respect to fuel efficiency, $\varepsilon_{F,E}$, is equal to the elasticity of the numerator of (1) minus that of the denominator (both with respect to E). The latter is one. The former, if fuel price is constant, is equal to the elasticity of VMT with respect to PM, which we denote as $\varepsilon_{M,PM}$, times the elasticity of PM with respect to E, which is −1. To summarize:

$$\varepsilon_{F,E} = -1 - \varepsilon_{M,PM} . \tag{2}$$

Since both $\varepsilon_{F,E}$ and $\varepsilon_{M,PM}$ are negative, the magnitude (i.e. absolute value) of $\varepsilon_{F,E}$ is smaller than one due to the effect of per-mile fuel cost on travel ($\varepsilon_{M,PM}$).

Equation (2) tells us how to decompose a change in fuel consumption into a direct effect and a rebound effect. Multiplying both sides by a fractional hypothetical change $\Delta E/E$ in fuel efficiency, (2) implies that fuel use declines as follows:

$$-\frac{\Delta F}{F} \equiv -\varepsilon_{F,E} \cdot \frac{\Delta E}{E} = \frac{\Delta E}{E} + \varepsilon_{M,PM} \cdot \frac{\Delta E}{E} . \tag{3}$$

The first term on the far right-hand side of (3), $\Delta E/E$, is the proportion that fuel consumption would be reduced in the absence of any changes in VMT. Let us call this the "direct impact" of the efficiency change on fuel consumption. The second term, which is negative, is the "take-back effect" from increased VMT: the reduction is less than the direct impact by this amount. We may define the rebound effect, a positive amount, as the negative of this take-back effect:

$$B = -\varepsilon_{M,PM} \cdot \frac{\Delta E}{E} \tag{4}$$

or, as a fraction of the direct effect,

$$b = -\varepsilon_{M,PM} . \tag{5}$$

The situation is depicted graphically in Figure 1. In the upper portion, curve DM represents the demand for VMT as a function of per-mile fuel cost, PM. The lower portion shows fuel consumption as a function of VMT at two different values for fuel efficiency: E^0 and E^1. Let efficiency increase from E^0 to E^1. If VMT remained constant at its initial value, M^0, fuel consumption would decline by amount ΔF_0 (the direct effect). But in fact VMT increases from M^0 to M^1, raising fuel consumption by amount R. The total reduction in fuel consumption is then not ΔF_0 but only $\Delta F = \Delta F_0 - R$. (In fact, there is no guarantee that ΔF is even positive; theoretically, it could represent a rise in consumption, in which case the rebound effect would be bigger than the direct effect.) The rebound effect is R in absolute terms, or $R/\Delta F_0$ in relative terms.

Figure 1. The rebound effect for an exogenous change in fuel efficiency
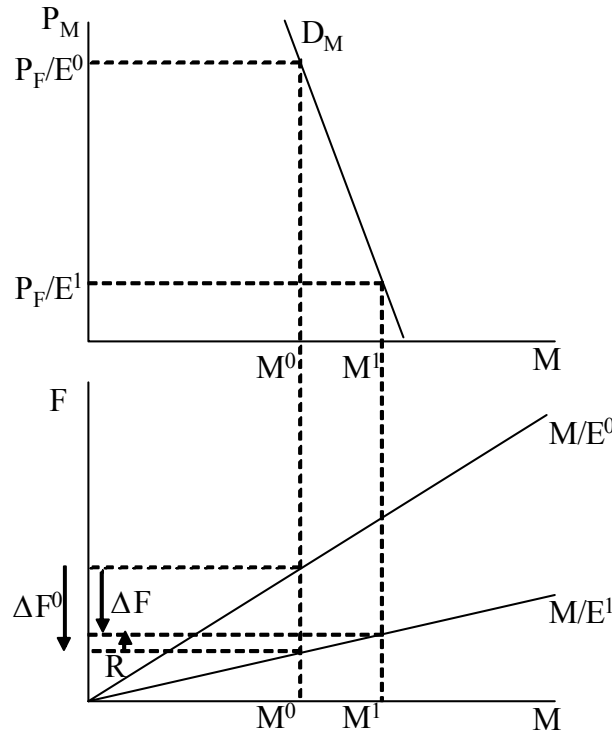
A problem with the theory in equations (1)–(3) is that fuel efficiency E is actually chosen by users, along with M, rather than being set by purely technological or regulatory considerations. This fact has been considered by US Department of Energy (1996, p. 5-11) in interpreting the empirical measurements of the rebound effect. Most of these measurements have taken advantage of changes in fuel price PF more than changes in efficiency E. Therefore, an obvious question is: what connection is there between the fuel-cost-per-mile elasticity of travel, $\varepsilon M, PM$, and the price elasticity of demand for fuel, $\varepsilon F, PF$? The answer can be determined from equation (1) by writing E as a function E(PF) of fuel price, with elasticity $\varepsilon E, PF$:

$$F\left(P_F, E\right) = \frac{\hat{M}\left[\dfrac{P_F}{E\left(P_F\right)}\right]}{E\left(P_F\right)}. \tag{1'}$$

Equation (1′) shows that $\varepsilon F, PF$ can be decomposed into two parts: the elasticity of $\hat{M}$ with respect to PF minus the elasticity of E with respect to PF. The second part is $\varepsilon E, PF$ by definition. The first part is found using the chain rule for differentiation on the numerator of (1′): it is just $\varepsilon M, PM$ times the elasticity of (PF/E) with respect to PF, which is one minus $\varepsilon E, PF$. Collecting these results:

$$\varepsilon_{F,PF} = \varepsilon_{M,PM} \cdot \left(1 - \varepsilon_{E,PF}\right) - \varepsilon_{E,PF}. \tag{6}$$

This equation also appears in USDOE (1996, p. 5-11). It provides an alternate way to measure the rebound effect, without ever measuring VMT elasticities directly. Rather, one measures the elasticities of fuel consumption and fuel efficiency, both with respect to fuel price; one then solves equation (6) for $\varepsilon_{M,PM}$, which as already explained is the proportional rebound effect:

$$\varepsilon_{M,PM} = \frac{\varepsilon_{F,PF} + \varepsilon_{E,PF}}{1 - \varepsilon_{E,PF}} \, . \tag{7}$$

To the best of our knowledge no one has actually estimated the rebound effect this way, but USDOE (1996, Table 5-2) uses (6) to relate a set of preferred elasticity estimates to each other.[2] Also, equation (6) makes it very clear that since empirical estimates of $\varepsilon_{F,PF}$ and $\varepsilon_{M,PM}$ differ greatly,[3] it must be that $\varepsilon_{E,PF}$ is considerably different from zero; this means we should not ignore the dependence of E on PF in using (1′) to guide empirical estimates.

Once we recognize that E depends on PF, we must consider more closely the determinants of that relationship. When we do so, we realize that E is chosen by consumers and/or producers. This has deeper implications, which make the very meanings of the elasticities in (2) ambiguous. This is because any reasonable theory of how E responds to fuel price would recognize that the answer depends on how much people travel. Consumers are unlikely to spend extra money on light-weight materials and special drive trains, or to give up desired features such as engine power and luggage space, in order to improve fuel efficiency for a little-used vehicle; but they may well do so for a vehicle they expect to drive a lot. Thus E depends on M. But the logic behind the elasticity $\varepsilon_{M,PM}$ in equation (2) is then circular: M depends on PM, which depends on E, which in turn depends on M. More generally, policies such as those aimed at carbon dioxide emissions, affect a complex set of decisions by manufacturers, such as model design and pricing, and by consumers, such as which models to purchase and how much to drive each one. Together these decisions influence all the quantities involved in (2). It is these kinds of complexities that we attempt to resolve in the Section 3.

Meanwhile, we can see right away that ignoring the dependence of E on PF causes the rebound effect to be overestimated. According to the argument just made, any unobserved factors that cause M to be large (perhaps an unusually long commute) also cause E to be large (as the commuter chooses vehicles to reduce the fuel cost of that long commute). This creates a negative correlation between M and PM≡PF/E, in addition to that which is in the demand function for vehicle-miles, M(PM). Thus the extent of the negative effect of PM on M will be overestimated. Equivalently, the magnitude of $-\varepsilon_{M,PM}$, the rebound effect, is overestimated.

---

[2] There is a typographical error in the last line of USDOE (1996, Table 5-2): "recent estimates" of the "fuel price elasticity of fuel economy" should read +0.200 rather than +0.220, as is clear from the text at the top of page 5-14 and also is necessary for consistency with USDOE's equation (2) on p. 5-11, which is the same as our equation (6).

[3] See USDOE (1996, pp. 5-14 and 5-83 to 5-87); Graham and Glaister (2002, p. 17); and the review in Parry and Small (2002, pp. 22-23).

2.1.2 Extensions of the Standard Definition

Besides the difficulties created by the endogeneity of E, there are several other simplifications in the standard definition just given that need to be relaxed in a more general treatment. First, fuel cost is just one of several components of the cost of using motor vehicles. One of the most important is time costs, i.e. the opportunity cost of the time spent by the occupants making their trips. Since that cost probably grows with rising incomes, it is quite likely that it becomes a larger portion of the cost of using vehicles over time in a growing economy. This would make the elasticity of VMT-demand with respect to fuel cost diminish over time, as fuel costs become relatively less important (Greene, 1992). For example, the cost of operating trucks is dominated by driver wages, making fuel relatively unimportant in decisions by firms about how to use their vehicles and in decisions by shippers about quantity and mode of shipments. More broadly, the robust growth in VMT over the last half century is presumably due in part to the dominance of time and convenience in people's travel decisions, suggesting that changes in fuel costs are likely to play only a minor role. Although for simplicity our theoretical section does not incorporate time trends or rising incomes, it is easy to do so in the empirical work we propose by specifying the empirical equation so as to allow the relevant elasticity to change with income or over time.

A second extension, also related to time costs, is to recognize that traffic congestion may be affected by the VMT changes that create the rebound effect. This effect may or may not be included in empirical measurements of M(PM), depending on the specification used. In any case, one may want to estimate explicitly whether and how much congestion is increased as a result of rebound effects. If congestion is substantially increased, the rebound effect would be diminished (because the tendency to increase VMT will be dampened by any increase in congestion), yet even a smaller rebound may be of greater concern due to the costly nature of congestion. Of course, any congestion increase tends to be somewhat self-limiting because congestion itself is a deterrent to more travel.

A third extension would be to consider that the price of fuel might be affected by policies affecting fuel demand, especially in a state as large and geographically isolated as California. Suppliers of California specific fuel mixes appear to be operating at or near capacity (USDOE, 2003). Furthermore, capacity may to some extent be strategically determined by producers who have some influence over price. Thus a reduction in fuel demand following an improvement in the fuel efficiency of cars may reduce the price of fuel. This tendency would make the rebound effect stronger: not only does a vehicle use less fuel per mile, but fuel itself also becomes cheaper.

2.2   Measuring the rebound effect

This subsection considers studies that estimate the function M(PM) in the numerator of equation (1) or (1′). Such studies attempt to explain vehicle-miles traveled on the basis of a number of variables, including the fuel cost per mile, PM or some proxy for it. As described in the previous subsection, the rebound effect can be defined in fractional terms as the negative of the elasticity of this function.

It is convenient to divide studies into those using aggregate data and those using micro (individual) data. The former includes both pure time-series studies (one observation per year) and aggregate cross-sectional time series studies which combine cross-sectional and time series variation (e.g. one observation per state per year).

2.2.1 Aggregate Time Series Studies

One group of studies takes advantage of changes in fuel prices and fuel efficiency over time to measure the effect of PM on vehicle-miles traveled. One of the most carefully done is that by Greene (1992), using US time series data for 1957-1989. Greene considers two complicating factors. First, some or all of the variables may affect VMT slowly, through lagged effects, if individuals respond only slowly to changed circumstances. Second, the unobserved influences on VMT that are captured by the error terms may be autocorrelated, i.e. errors for observations close together in time may tend to move together; this could be caused by unobserved factors that persist over time, such as lifestyle factors, or by errors in data measurement that are similar from one year to the next.

Greene finds the rebound effect over the entire time period 1957-89 to be between 5 and 15 percent both in the short and long run, with a best estimate of 12.7 percent. Lagged effects are found not to be important, which is why the long-run effect is the same as the short-run.[4] He also finds it is important to account for autocorrelation, which is quite strong, the correlation between errors in adjacent years being estimated at 0.74. Failing to account for autocorrelation results in spurious measurements of lagged values; Greene believes this has misled other researchers to find long-run effects to be spuriously greater than short-run effects.[5]

Finally, Greene finds some evidence that the fuel-cost-per-mile elasticity declines over time, thus tentatively confirming the first possibility mentioned in Section 2.1.2. There are two lines of evidence. First, Greene notes that in a linear model the elasticity necessarily declines over time because P_M has decreased and VMT increased over most of the period; and he finds little difference in quality of results between a linear and a non-linear specification. Second, in the double-log model that Greene presents for most of his results, in which the elasticity itself is an estimated parameter, he tries specifications in which this elasticity (and hence the rebound effect) is allowed to differ across time periods; these results show substantial declines in the rebound effect, from 27.4 percent in 1966-77 to 5.9 percent in 1978-89. These declines are statistically significant at the 10 percent but not at the more rigorous 5 percent significance level. Given the sample of only 24 years forming the basis for this test, it is not surprising that it is only

---

[4] However, Greene's explanation demonstrates the difficulty of interpreting equations in which fuel cost per mile (P_F/E) is a variable, when in fact E is exogenous. Greene notes that in the long run consumers react to an increase in fuel price in part by increasing fuel efficiency. But rather than reducing the long-run effects, as he argues, this ideally would call for a specification that permits fuel efficiency to react with a lag to changes in fuel price.

[5] Another study that found autocorrelation is that by Blair, Kaserman, and Tepel (1984). They obtain a rebound effect of 30 percent, based on monthly data from Florida from 1967 through 1976. They did not estimate models with lagged variables.

of marginal statistical significance, and we regard the very large decline in magnitude as moderately strong evidence for a shift over time.

Jones (1993) re-examines Greene's data, after including observations for 1990, focusing on model selection issues in time series analysis. Jones confirms that the autoregressive model selected by Greene is statistically valid. However, alternative specifications, notably those that include lagged dependent variables, are acceptable as well. Such models do produce long-run estimates of the rebound effect of about 31%, exceeding the short-run estimates of ca. 11%.[6] In Jones' own interpretation, the results confirm Greene's estimates of the short-run rebound effect but weaken Greene's conclusion regarding the absence of a larger long run effect.

Schimek (1996) uses data from a longer time period than Greene (1992) and finds a similarly small or even smaller short-run rebound effect. But he obtains a larger long-run rebound effect, about 30 percent, similar to Jones (1993). In Schimek's preferred results, the short-run and long-run rebound estimates are 7 and 29 percent.[7] He accounts for federal Corporate Average Fuel Economy (CAFE) regulations by including a time trend for years since 1978, and he also includes dummy variables for the years 1974 and 1979 when gasoline rationing was in effect. Schimek finds that these controls improve the estimates by reducing signs of autocorrelation (Durbin's h-statistic falls from 2.60 to 0.93).[8]

Schimek (1996, Table 2) also estimates three equations which decompose fuel consumption into vehicle stock, fleet-average fuel efficiency, and driving per vehicle. All are in double-log form. The third of these equations permits fuel price and fuel efficiency to have distinct effects; and the estimated coefficients are opposite in sign and nearly identical in magnitude, as would be expected if they enter as a ratio as assumed in the specification used by most authors. This equation also suggests a rather small short-run rebound effect (5–6 percent) and a much larger long-run effect (21–26 percent). The three-equation framework offers an ideal opportunity to explore the empirical consequences of simultaneity between the determinants of VMT and fuel efficiency, but Schimek does not explore this aspect.

The possible significance of lagged dependent variables is very important for sorting out short-run and long-run effects. However it is unsettled in the pure time-series studies, because it is difficult to disentangle the presence of a lagged dependent variable from the presence of autocorrelation. Greene finds that when autocorrelation is accounted for, lagged dependent variables are not longer statistically significant. Schimek finds that when lagged dependent variables are included, evidence for autocorrelation becomes weak. Neither of these results demonstrates definitively which is the right specification, and the answer appears sensitive to the time period considered and the way in which CAFE standards are controlled for. Fortunately, cross-sectional time series studies offer an opportunity to obtain a larger data set with the hope of disentangling these two statistical properties.

---

[6] Estimate from the linear lagged dependent variable model (model III in Table 1). Estimates for the loglinear model are nearly identical.
[7] Schimek (1996), p. 87, Table 3, model (3).

[8] The CAFE variable makes even more difference in another equation, explaining fuel consumption, where without the CAFE variable income has the wrong sign and the lagged dependent variable takes an unreasonably large coefficient. See his Table 1, models (1) and (4).

2.2.2 Aggregate Cross-sectional time series Studies

Haughton and Sarkar (1996) construct a much larger data set which allows them to take advantage of differences not only across time but across states. Because fuel prices vary by state, primarily due to different rates of fuel tax, this provides additional opportunity to observe its effects on amount of motor-vehicle travel. This cross-sectional time series data set for the US covers years 1970-1991 and all of the 50 states plus the District of Columbia. Furthermore, the authors estimate both VMT per driver and, in a separate equation, fuel intensity (the inverse of fuel efficiency). Unfortunately, however, the simultaneity of these two equations is not accounted for.

Haughton and Sarkar's estimate of the rebound effect is about 16 percent in the short run, similar to Greene (1992), and 22 to 23 percent in the long run.[9] In obtaining long-run estimates, they are able to measure both autocorrelation and the effects of a lagged dependent variable with sufficient precision to resolve some of the difficulties with pure time series. For autocorrelation, they measure the correlation between adjacent years to be 0.38–0.48. Like Greene, they find that when autocorrelation is accounted for, the effect of lagged dependent variable is much diminished; but unlike Greene they still obtain a statistically significant effect, one that implies a long-run effect 32 to 45 percent larger than the short-run effect.

As for fuel intensity, Haughton and Sarkar find it is unaffected by the current price of gasoline unless that price exceeds its historical peak. This phenomenon is known as hysteresis, and is rationalized on the grounds that consumers and manufacturers react to increases in fuel price by investing in new technologies which, once adopted, are continued even if price declines thereafter. CAFE effects are taken into account through a variable measuring the difference between the legal minimum and actual fuel efficiency in 1975; however, that variable is so strongly correlated with the historical maximum real price of gasoline that they omit it in most specifications, casting doubt on whether the resulting estimates really control adequately for the CAFE regulation.

It appears that the confounding of the rebound effect with effects of CAFE regulation is one of the most limiting factors in studies reviewed so far. Different authors have defined and included a variety of variables to account for this, and results seem sensitive to just how it is done (Schimek 1996). Probably this is because the standards were imposed about the same time as a major increase in fuel prices occurred, and they became more stringent as incomes rose during the 1980s; therefore the effects of CAFE standards are hard to separate from those of fuel prices and incomes. However, the problem is compounded by the fact that there has been little agreement or theoretical justification for the particular forms used. With results so sensitive to inclusion of one of these variables, we give a high priority to developing a theory to provide better guidance for specifying the demand equation so as to account for such regulations. In conjunction with accounting for simultaneity of fuel efficiency and travel, it should be possible to obtain more reliable results.

---

[9] This paragraph is based on models E, and F in their Table 1, p. 115. Their variable, "real price of gasoline per mile," is evidently the same as fuel cost per mile.

2.2.3 Disaggregate Studies

Recently some studies have used data on individuals, either a cross section in a single year or from a panel covering several years' observations on the same people. In reviewing such studies, Greene, Kahn, and Gibson (1999) note that the disparity of results for the rebound effect is wider than for aggregate studies, covering a range from zero to about 50 percent.

Goldberg (1998) uses a variety of sources, including the Consumer Expenditure Survey, to estimate a suite of models explaining manufacturers' design and pricing decisions as well as consumers' vehicle purchase and usage decisions. One important innovation is the use of instrumental variables to account for simultaneity between the vehicle purchase and vehicle usage decision. This innovation makes a big difference to the estimated rebound effect, reducing it from about 20 percent to essentially zero. However, in the usage equation with instrumental variables, the variables representing vehicle type attain astronomical yet statistically insignificant coefficients (Goldberg's Table I), casting some doubt on the ability of the data set to measure this simultaneity and hence on the reliability of the zero-rebound result. Furthermore, the utilization equation is estimated using data only on households who purchased a new car the previous year, so is not necessarily representative of all vehicle users.

Pickrell and Schimek (1999) estimate a vehicle-use model with 1995 cross-sectional data from the National Personal Transportation Survey (NPTS). The elasticity of VMT with respect to gasoline price, controlling for ownership levels,[10] is –0.04 (model 3 with odometer readings as dependent variable). This low figure emerges when residential density is included as explanatory variable; residential density is collinear with the fuel price, so that it is hard to separate their effects. More generally, the value of a cross-sectional micro data set for a single year is diminished by the fact that fuel prices vary only across states, and those variations may be correlated with unobserved factors that also influence VMT.

Greene, Kahn, and Gibson (1999) use a series of large micro data sets covering six different years between 1979 and 1994. The data are drawn from the Residential Energy Consumption Survey in 1979 and 1981, and from Residential Transportation Energy Consumption Surveys (RTECS) performed approximately every three years from 1985 to 1994. The authors derive demand functions for VMT and for fuel within a household production model. This leads to separate equations explaining VMT, fuel price,[11] and fuel efficiency, for each owned vehicle and for households with each of four different ownership levels of passenger vehicles (one, two, three, and four or more vehicles). Vehicle ownership is assumed exogenous. In order to account for CAFE regulations, the usage equations include as an explanatory variable the average fuel efficiency of all cars produced in the same model year as the vehicle whose use is being explained.

---

[10] Hence for fuel efficiency, at least to a large extent.

[11] The fuel price is endogenous because as drivers can react to higher fuel prices by searching harder for cheaper fuel. Also, people who drive longer distances are likely to incur lower search costs for cheap fuel, as they simply pass more gas stations.

The rationale for letting households choose fuel efficiency, even for a fixed vehicle fleet, is that changes in fuel prices may lead households to adapt the way they use their vehicle portfolio in a way that affects average fuel efficiency. A particularly interesting example is that households might respond to lower fuel cost per mile by taking longer trips (e.g. the occasional holiday trip); this would tend to increase gasoline consumption (a type of rebound effect) yet could actually increase average fuel efficiency since long trips are more fuel-efficient than short ones.

Two results in Greene et al. are noteworthy. First, the authors test whether consumers care separately about fuel price PF and fuel efficiency E, as opposed to their ratio PM, in their usage decision. This is a formal test of the equal and opposite signed coefficients already noted from Schimek's study, and it confirms that entering PF and E as a ratio is valid. Second, the long-run rebound effect is estimated at 23 percent overall, with a range from 17 percent for three-vehicle households to 28 percent for one-vehicle households.

2.3   Assessing the effectiveness of the CAFE regulation

This section discusses some studies that do not focus on estimating the rebound effect, but which are of interest because they present frameworks for analyzing CAFE type of regulations and produce some provocative results.

Kleit (1990) builds an aggregate simulation model, using data and elasticities from secondary sources, to analyze the effects of stricter CAFE standards. He shows that the CAFE standard may increase the sales of new cars, because the standard acts like an implicit tax on large cars and an implicit subsidy on small cars. Consequently, when supply and/or demand elasticities in the small car markets are relatively large, total new car sales increase after introduction of the standard. Kleit then assumes that the reduction in average driving costs after the change in the car mix (towards small cars) leads to increased driving because of the rebound effect.[12] Consequently, gasoline consumption may increase with stricter CAFE standards. However, the actual direction the effects take depends on the stringency of the standard and on the parameters of the demand system.

Kleit (2000) updates and extends Kleit (1990). The newer study allows for technology-forcing standards as well as induced shifts in the mix of vehicles. It is found that making CAFE stricter will increase VMT, and therefore will increase the emissions of traditional pollutants (but not of CO_2).[13] This of course is an inherent problem with any VMT increase when pollutants are regulated using per-mile emissions standards; but actually most of the increased emissions

---

[12] This is a fairly loose interpretation of the rebound effect, as in Kleit's analysis vehicle characteristics apart from fuel efficiency are not constant. But when consumers are forced to buy smaller cars, they buy a less preferred car. Driving that car generates less utility than the more preferred car, and this reduces car use. That reduction at least partially compensates the effects of the lower per mile cost, so that Kleit's estimates of VMT impacts should be considered to be upper bounds.

[13] Kleit also questions if further reducing gasoline consumption is beneficial, as it is already highly taxed. Parry and Small (2003) argue that the present US gasoline taxes are too low, however.

predicted by Kleit's model are due not to VMT but rather to the shift to light trucks caused by their lower energy efficiency requirement under federal CAFE regulations.

In a similar vein, Thorpe (1997) uses a general equilibrium model of the new car market to assess the extent to which a shift in vehicle mix dilutes CAFE's ability to improve average fuel efficiency. Using his central scenario parameter and elasticity values,[14] he finds that average fuel efficiency actually decreased because of CAFE. The reason is that in his model, Asian and European producers reduce their sales of small fuel-efficient cars, increasing their sales of midsize cars (attracting former buyers of US-built midsize cars). Furthermore, all producers increase their truck sales. The overall mix shift hence is toward relatively less efficient cars. The insight is that the average efficiency of new sales is attracted to the average fuel efficiency standard for all sellers, both those that did and did not exceed the standard before its implementation.

Goldberg (1996) performs a comprehensive structural analysis of fuel-economy regulations. She constructs a disaggregate discrete/continuous choice model, estimates it, and applies it to analyze policies to reduce fuel consumption. The main finding is that CAFE has worked well, and that fuel taxes would need to be very high to produce similar results. This insight depends crucially on her econometric results, already discussed, which include many crucial coefficients that are statistically insignificant.

Finally, one could take a broader perspective and analyze the effect of CAFE regulations on the transport sector or the economy as a whole, rather than just fuel consumption. In doing this, the interaction between different transport externalities should be taken into account in assessing policies addressing one particular externality, like global warming (Harrington and McConnell, 2003). This leads to a number of observations. For example, since congestion is believed by most researchers to be the most important externality quantitatively, an increase in VMT (possibly due to the rebound effect) may be very costly to society. Similarly, the combination of increased VMT and fixed per-mile emission standards for many pollutants will lead to increased emissions of these pollutants, with potentially serious damage costs. There has been widespread concern that stringent fuel efficiency standards could lead to less safe vehicles, although the evidence is mixed and the effects apparently complex (NRC, 2002). Finally, it can be argued that pollution-abatement policies that raise revenues (like taxes) are to be preferred over those that do not (like standards) due to their effects on the efficiency of the overall tax system (Parry and Oates, 2000).

2.4   Conclusions from the Literature Survey

In summary, this section shows that aggregate estimates of the short-run rebound effect are fairly robust. Estimates of the long-run rebound effect, by contrast, are sensitive to the particular specification, especially the treatment of time patterns and CAFE standards.

---

[14] The central scenario produces the strongest decrease in average fuel efficiency among all alternative scenarios presented in Table 5. It is not the scenario with the average parameter values, and alternative scenarios increase average fuel efficiency.

Disaggregate cross-sectional analyses tend to produce a greater range of estimates. One disaggregate study that exploits both cross-sectional and temporal variation (Greene et al., 1999) finds a long-run rebound effect of 23 percent, similar to several other studies. Finally, some analyses of the CAFE regulation show that this type of fuel efficiency regulation can have – and indeed has had – unexpected consequences.

In the following sections we take account of these lessons from the literature review:

- The standard definition of the rebound effect refers to exogenous changes in fuel efficiency, but it is measured mostly by exploiting variations in fuel price. In reality fuel efficiency is an endogenous variable, even in the presence of fuel-efficiency regulation.

- A reliable estimate of the rebound effect from US data requires accounting for the effects of federal CAFE regulations.

- Empirical estimates of the rebound effect show considerable sensitivity to the specification, especially of the time-series properties of the data (autocorrelation and lagged variables). It appears that combined cross-sectional and time-series observations offer the best way to observe sufficient numbers of observations to control for these aspects of the data.

The theoretical model in the next section provides a precise way to incorporate the first two refinements. The third point is addressed in our empirical approach.

3. Theoretical Refinements

In this section we formulate a theoretical model to provide a guideline for the econometric estimation of the rebound effect using aggregate data, taking account of some concerns raised in the previous section. An important difference with the standard approach is that we introduce an explicit relation between a regulatory variable and fuel efficiency, allowing us to treat fuel efficiency as an endogenous variable. This also permits a flexible but well-defined description of how the model system is influenced by regulations whose direct or indirect effects are to increase fuel efficiency.

3.1 Structural-Form and Reduced-Form Models

We assume that consumers choose how much to travel (M) on the basis of their vehicle ownership, the per-mile cost of driving and exogenous characteristics. They choose how many vehicles to own (V) on the basis of the price of new vehicles, how much they intend to drive, the cost of driving, and other characteristics. The fuel efficiency choice (E) is determined jointly by consumers and manufacturers taking into account the price of fuel, how much they intend to drive, the regulatory environment, and other characteristics.

These assumptions constitute a *structural model*, in the sense that they lay out the decision-making structure explicitly. Mathematically, these assumptions are characterized in the following set of equations:

$$M = M\left(P_M, V, X_M\right)$$
$$V = V\left(P_V, M, P_M, X_V\right) \tag{8}$$
$$E = E\left(P_F, M, R_E, X_E\right)$$

where M is aggregate VMT; V is the size of the vehicle stock (not its composition); P_V is a price index for the ownership cost of new vehicles; P_F is a price index for fuel; $P_M \equiv P_F/E$ is the fuel cost per mile; X_M, X_V and X_E are exogenous variables affecting M, V and E, respectively; and R_E represents one or more regulatory variables. Vehicle price P_V represents the user cost of capital for a vehicle purchase, which incorporates not only the price itself but financing or leasing costs. The regulatory variables could represent any of a wide variety of measures that directly or indirectly influence fleet-average fuel efficiency.

The endogenous variables in system (8) are those on the left-hand side. All three also appear on the right-hand side (since E is part of the definition of P_M). We make this explicit as follows, writing the endogenous variables in bold type for clarity:

$$\boldsymbol{M} = M\left(\frac{P_F}{\boldsymbol{E}}, \boldsymbol{V}, X_M\right)$$
$$\boldsymbol{V} = V\left(P_V, \boldsymbol{M}, \frac{P_F}{\boldsymbol{E}}, X_V\right) \tag{9}$$
$$\boldsymbol{E} = E\left(P_F, \boldsymbol{M}, R_E, X_E\right)$$

Equation set (9) constitutes three equations in three unknowns (M, V, and E). If we solve it for those unknowns, we can write each of them in terms of all the other variables, i.e. the exogenous variables. This way of writing the equations makes clear how each variable depends on things that can be independently varied, and so is useful for empirical work. It is known as the *reduced form* of the equations, denoted here by the symbol ~:

$$M \equiv \tilde{M}\left(P_V, P_F, R_E, X_M, X_E, X_V\right)$$
$$V \equiv \tilde{V}\left(P_V, P_F, R_E, X_M, X_E, X_V\right) \tag{10}$$
$$E \equiv \tilde{E}\left(P_V, P_F, R_E, X_M, X_E, X_V\right)$$

The first of these equations shows explicitly, for example, that vehicle prices (and other determinants of the ownership cost of capital invested in new vehicles) do affect VMT, even though they do not affect operating cost as depicted in the first of equations (9). Their effect is indirect, via the second of equations (9). Presumably a higher vehicle price causes vehicle purchases to decline (lower V), which then calls for VMT to decline (lower M). There could also be an income effect, not accounted for in the model as written, resulting from changes in expenditures on purchasing vehicles; but this is likely to be smaller than the effect portrayed here and in fact is not predictable in sign because raising the price of vehicles could cause

expenditures on vehicles either to rise or to decline, depending on the price elasticity of the demand function V(PV,PM,XE). In any case, income effects will be captured by the elasticities estimated in the empirical work.

All of these equations, especially the second and third, are probably autoregressive in nature, given that vehicle stock in a given year will be heavily influenced by the stock in the previous year and similarly some determinants of usage, like housing and job location, cannot be changed quickly. Although for simplicity our theoretical notation does not show this autoregressive structure, the empirical specification will do so.

To understand some of the empirical literature, it is useful to define a partially reduced form for the usage equation, denoted here by ^. The reason for doing this is to clarify what is being estimated in usage studies that examine VMT as a function of PM but not of other endogenous variables. In this form the second of equations (8) is substituted into the first, while leaving both still as functions of the endogenous variable PM:

$$M = M\big[P_M, V(P_V, M, P_M, X_V), X_M\big] \tag{11a}$$

which can be solved for M as a function of the other variables:

$$M = \hat{M}(P_M, P_V, X_M X_V,). \tag{11b}$$

This equation corresponds to an empirical equation for usage in which vehicle stock is not included and in which efficiency is included only indirectly via the per-mile fuel-cost variable. Equation (11b) shows that such an equation should include the exogenous variables that influence the vehicle stock, PV and XV; their influence on M arises through their influence on V, as seen explicitly in (11a). If PV, XV, and XM are all held constant, equation (11a,b) is the same as our earlier definition of function $\hat{M}(\cdot)$ in (1). Therefore the elasticity identified as the rebound effect in (4) and (5) can be written more generally as the elasticity of this function:

$$\varepsilon_{\hat{M},PM} \equiv \frac{P_M}{M} \cdot \frac{\partial \hat{M}}{\partial P_M}. \tag{12a}$$

But the rebound effect can also be written in terms of the structural equations (9), which enables one to see explicitly how much of it comes from changes in usage per vehicle and how much in changes in number of vehicles and their subsequent effect on usage. This is done by applying the chain rule for differentiating (11a), at its solution as given by (11b), then multiplying by (PM/M) to convert to elasticities, and solving for $\varepsilon_{\hat{M},PM}$. The result is:[15]

---

[15] Since (11b) is the solution to (11a), we can write (11a) as follows, omitting for clarity the variables P_V and X_V, which are fixed in this calculation):

$$\hat{M}(P_M) = M\Big[P_M, V\big(P_M, \hat{M}(P_M)\big)\Big].$$

Differentiating,

$$\varepsilon_{\hat{M},PM} = \frac{\varepsilon_{M,PM} + \varepsilon_{M,V}\varepsilon_{V,PM}}{1 - \varepsilon_{M,V}\varepsilon_{V,M}} \tag{12b}$$

where εM,PM and εM,V and elasticities of the first of equations (9) (with respect to PM and V, respectively) and εV,M is the elasticity of the second (with respect to M).

To summarize we have discovered two features of a valid empirical specification for measuring a partially reduced form relating travel M to the per-mile fuel cost P_M. (Such a function lies at the heart of most empirical measurements of the rebound effect, but not ours.) First, the empirical specification should include as independent variables all the exogenous factors determining both vehicle stock and usage but not vehicle stock itself, which has been substituted out in deriving (11a,b). Second, the equation needs to be estimated using a technique to account for the endogeneity of E in forming variable P_M. A common such technique is instrumental variables, which uses as instruments a set of variables that are expected to influence usage indirectly through the endogenous variable E but not directly. A comparison of the variables in the last of equations (10) with those in (11a,b) tells us immediately what variables to use as instruments: P_F, R_E, and X_E. That is, the instruments should include price of fuel, descriptors of regulation, and other exogenous variables influencing fuel efficiency but not directly influencing vehicle stock or usage. This is different from the approach taken in most of the empirical literature, which has placed regulatory descriptors directly in the usage equation and, except in a few cases, not used instrumental variables to account for endogeneity of P_M.

3.2 The rebound effect

Extending our earlier definition, the rebound effect can be defined as the difference between the reduction in fuel consumption after some regulatory change, $-\Delta F$, and the reduction that would occur under some simplified calculation, which we can call the "direct impact," $-\Delta F_0$. As before, we define it to be positive and as a fraction either of initial fuel consumption or of the direct effect:

$$B = \frac{\Delta F - \Delta F_0}{F}$$
$$b = \frac{\Delta F - \Delta F_0}{-\Delta F_0}. \tag{13}$$

$$\frac{\partial \hat{M}}{\partial P_M} = \frac{\partial M}{\partial P_M} + \frac{\partial M}{\partial V}\left(\frac{\partial V}{\partial P_M} + \frac{\partial V}{\partial M}\frac{\partial \hat{M}}{\partial P_M}\right)$$

or, in elasticity terms:

$$\varepsilon_{\hat{M},PM} = \varepsilon_{M,PM} + \varepsilon_{M,V}\left(\varepsilon_{V,PM} + \varepsilon_{V,M}\varepsilon_{\hat{M},PM}\right).$$

Solving this equation for $\varepsilon_{\hat{M},PM}$ gives (12b).

Our earlier definition corresponds to the case where the direct impact is calculated as though any change in fuel efficiency, ΔE, translates directly to fuel savings with travel kept constant:

$$\Delta F_0 = M \cdot \Delta\left(\frac{1}{E}\right) = -\frac{M}{E^2} \cdot \Delta E = -F \cdot \frac{\Delta E}{E} \tag{14}$$

The true change in F is

$$\Delta F = \Delta\left(\frac{M}{E}\right) = \Delta F_0 + \frac{\Delta M}{E} . \tag{15}$$

Defining this version of the rebound effect as the "basic" rebound, B_0 or b_0, it is:

$$B_0 = \frac{1}{F} \cdot \frac{\Delta M}{E} = \frac{\Delta M}{M}$$
$$b_0 = \frac{\Delta M / M}{\Delta E / E} = -\varepsilon_{M,PM} . \tag{16}$$

The reduced form equation for E in (10) makes clear that the fuel efficiency change driving these definitions can take place through several different routes. We will consider three: a change in the regulatory parameter R_E, a change in fuel price P_F, and a change in vehicle prices P_V. Because these three exogenous variables all have different impacts on E via the other equations, the effects of these on other variables like VMT will not be the same. For simplicity we treat each possible route separately, though of course all three might occur simultaneously from any given policy.

3.2.1 Rebound Effect from Regulatory Change

Suppose first there is a change in the regulatory parameter ΔR_E. According to the last of equations (10), this will cause some total change in fuel efficiency given by

$$\Delta E = \frac{\partial \tilde{E}}{\partial R_E} \cdot \Delta R_E . \tag{17}$$

However the reduced form equation is not necessarily what is most easily measured and certainly not what is most easily conceptualized, since the change ΔE in this equation takes place through complex adjustments of the entire system of structural equations (8). Rather, researchers are more likely to know something about the parameters of these structural equations (8). In fact, it is quite natural that the change in fuel efficiency from a given regulation would be initially predicted not as (17) but from the way the structural equation (8) depicts the effect of R_E on E:

$$\Delta E_1 \equiv \frac{\partial E}{\partial R_E} \cdot \Delta R_E . \tag{18}$$

17

We may call this the *direct change in fuel efficiency caused by the regulatory change*. We can use system (8) and the definition of P_M to find the changes in all three endogenous variables (V, M, and E) by applying chain rule differentiation:

$$\Delta M = \frac{\partial M}{\partial V} \cdot \Delta V + \frac{\partial M}{\partial P_M} \cdot \left(-\frac{P_F}{E^2} \right) \cdot \Delta E$$

$$\Delta V = \frac{\partial V}{\partial P_M} \cdot \left(-\frac{P_F}{E^2} \right) \cdot \Delta E \qquad (19)$$

$$\Delta E = \Delta E_1 + \frac{\partial E}{\partial M} \cdot \Delta M.$$

These three equations can be solved for the three unknowns ΔM, ΔV, and ΔE in two steps. First, substitute the second into the first to obtain:

$$\Delta M = M \cdot \left(-\beta_{M,PM} - \beta_{M,V} \cdot \beta_{V,PM} \right) \cdot \Delta E / E \qquad (20)$$

where each β represents an elasticity of the structural equations, for example $\beta_{M,V} \equiv (V/M) \cdot (\partial M / \partial V)$. The system is then solved by substituting (20) into the third of equations (19). Before doing so, we can simplify notation by recognizing that we already have a name for the quantity in parentheses in (20): namely, the basic rebound effect, b_0, as seen by comparing (20) with (16). That is,

$$-\varepsilon_{M,PM} = b_0 = -\beta_{M,PM} - \beta_{M,V} \cdot \beta_{V,PM} . \qquad (21)$$

Each of the terms in (21) is positive (since two of the three elasticities are negative); so we see right away that the basic rebound effect b_0 would be underestimated if we tried to measure it using the corresponding elasticity $-\beta_{M,PM}$ from the structural equation instead of using the partially reduced form equation (11).

Carrying out the last substitution leads to an equation we can easily solve for ΔE, obtaining:

$$\Delta E = \frac{\Delta E_1}{1 - b_0 \cdot \beta_{E,M}} \equiv a_1 \cdot \Delta E_1 \qquad (22)$$

where $\beta_{E,M}$ is the elasticity of fuel efficiency with respect to mileage in the structural equation for fuel efficiency, and a_1 is just a convenient notation for the reciprocal of the denominator. If, as seems most likely, high VMT encourages more fuel-efficient cars, so that $\beta_{E,M} > 0$, and if neither b_0 nor $\beta_{E,M}$ is very large, we can expect that $\Delta E > \Delta E_1$. However high VMT could encourage a desire for greater performance involving less fuel efficiency, in which case the inequality goes the other direction.

If ΔE_1 is considered the natural initial prediction of the effect of a regulation on E, then a natural definition of a "direct effect" of the regulation on fuel consumption is that which would occur given change ΔE_1 and no change in VMT: that is, ΔF_1 defined by

$$\Delta F_1 / F \equiv -\Delta E_1 / E \, . \tag{23}$$

This leads to an alternative definition of a rebound effect in analogy with (13):

$$b_1 \equiv \frac{\Delta F - \Delta F_1}{-\Delta F_1} = b_0 \cdot \frac{1 - \beta_{E,M}}{1 - b_0 \beta_{E,M}} \, . \tag{24}$$

Under the same conditions for which that $\Delta E > \Delta E_1$, this definition shows that $b_1 < b_0$. Under these conditions, this alternative rebound is based on a smaller definition of the "direct impact;" even though it is smaller in absolute amount than the earlier definition of rebound, it is larger as a proportion of the direct impact.

3.2.2 Rebound Effect from Change in Fuel Price

Now suppose there is instead a change in fuel price ΔP_F. Again according to the last of equations (10), this will cause some total change in fuel efficiency given by

$$\Delta E = \frac{\partial \widetilde{E}}{\partial P_F} \cdot \Delta P_F \, . \tag{25}$$

Once again, however, the reduced form equation is not necessarily the best for measuring or conceptualizing the changes in fuel consumption. Rather, we may want to turn to structural equation (8) to depict the effect of P_F on E, which we may define as:

$$\Delta E_2 \equiv \frac{\partial E}{\partial P_F} \cdot \Delta P_F \equiv E \cdot \beta_{E,PF} \cdot \left(\Delta P_F / P_F \right) \, . \tag{26}$$

We call this the *direct change in fuel efficiency caused by a change in fuel price*. The system of equations (19) now applies with the additional term $(\partial M / \partial P_M) \cdot (\Delta P_F / E)$ in the equation for ΔM, with the additional term $(\partial V / \partial P_M) \cdot (\Delta P_F / E)$ in the equation for ΔV, and with ΔE_2 substituted for ΔE_1 in the equation for ΔE. This modifies (20) as follows, where we have substituted b_0 for the term in parentheses there:

$$\Delta M = M \cdot b_0 \cdot \left(\frac{\Delta E}{E} - \frac{\Delta P_F}{P_F} \right) \, . \tag{27}$$

Equation (27) verifies that when estimating the determinants of VMT in a context where both efficiency and fuel price are changing, as is usually the case, the coefficient of the two should indeed be equal and opposite, as discussed in the literature review. Substituting this into the last of the modified equations (19), we can solve for ΔE to obtain:

$$\Delta E = \frac{1}{1 - b_0 \cdot \beta_{E,M}} \cdot \left[\Delta E_2 - E \cdot b_0 \cdot \beta_{E,M} \cdot \frac{\Delta P_F}{P_F} \right]$$

$$= \left[\frac{1 - \left(b_0 \cdot \beta_{E,M} / \beta_{E,PF} \right)}{1 - \left(b_0 \cdot \beta_{E,M} \right)} \right] \cdot \Delta E_2 \equiv a_2 \cdot \Delta E_2 \qquad (28)$$

where a_2 is defined as the term in square brackets in the lower equation. Assuming that $0 < b_0 \cdot \beta_{E,M} < \beta_{E,PF} < 1$, which seems likely provided higher VMT leads to choice of more fuel-efficient vehicles, we can see that $0 < a_2 < 1$, i.e. $\Delta E < \Delta E_2$. Once again, the true effect on fuel efficiency is smaller than one would predict just from the structural equation for E, due to feedback effects involving the rebound effect.

As in subsection 3.2.1, we can consider an alternative rebound effect derived on the assumption that the "direct impact" on fuel consumption, here denoted by $-\Delta F_2$, is calculated in analogy with equation (23) and the rebound effect is taken relative to ΔF_2:

$$\Delta F_2 / F \equiv -\Delta E_2 / E . \qquad (29)$$

$$b_2 \equiv \frac{\Delta F - \Delta F_2}{-\Delta F_2} = 1 - a_2 \left(1 - b_0 \right)$$

$$= b_0 \frac{1 - \beta_{E,M} + \left(1 - b_0 \right) \beta_{E,M} / \beta_{E,PF}}{1 - b_0 \beta_{E,M}} = b_1 + b_0 \frac{\left(1 - b_0 \right) \beta_{E,M} / \beta_{E,PF}}{1 - b_0 \beta_{E,M}} . \qquad (30)$$

Since b_0 and the fraction by which it is multiplied on the far right of (30) are positive, this definition shows that $b_2 > b_1$.

It is particularly interesting to compare a_1 with a_2 because doing so shows how a given direct change in fuel efficiency has different ultimate effects depending on whether it arises from regulation or fuel price. That is, if two situations produced the same direct change, $\Delta E_1 = \Delta E_2$, the actual change ΔE would be different. Specifically, it would be smaller for the case where the direct change arose from fuel prices than where it arose from regulation, provided $a_2 < a_1$ as is the case under the assumption that $\beta_{E,M} > 0$.

In order to numerically illustrate the results up to now, suppose the elasticity of mileage demand with respect to fuel cost per mile equals -0.2, the elasticity of fuel efficiency with respect to mileage is 0.1, and the elasticity of fuel efficiency with respect to the fuel price per gallon is 0.15. Plugging those values into equations (16), (24) and (30) produces a value for the standard rebound effect b_0 (defined as a proportion of the direct effect) of 0.2; the rebound effect from a regulatory change (b_1) equals 0.18; and the rebound effect from a fuel price change (b_2) equals 0.29.[16]

---

[16] Some special cases are interesting. When the elasticity of fuel efficiency with respect to fuel price equals one, then the standard rebound effect and the rebound effect from a fuel price change are the same. When the elasticity of fuel efficiency with respect to mileage goes to zero, then all rebound effect definitions are equivalent.

3.2.3 Rebound Effect from Change in Vehicle Prices

Finally, suppose there is a change in the average vehicle price ΔP_V. This could happen, for example, from technology-forcing regulatory standards. According to the reduced form equations (10), this will cause a total change in fuel efficiency given by

$$\Delta E = \frac{\partial \tilde{E}}{\partial P_V} \cdot \Delta P_V . \tag{31}$$

However, we again may want to turn to structural equation (8) to depict the effect of P_V on E, which we may define as:

$$\Delta E_3 \equiv \frac{\partial E}{\partial P_V} \cdot \Delta P_V = 0 . \tag{32}$$

In this case the *direct change in fuel efficiency caused by a change in the vehicle price* equals zero, because a vehicle price change, in contrast to a regulatory change or a fuel price change, has no direct effect on the demand for fuel efficiency.[17] The system of equations (19) now applies with the additional term $(\partial V / \partial P_V) \Delta P_V$ in the equation for ΔV, and with $\Delta E_3 = 0$ substituted for ΔE_1 in the equation for ΔE. This modifies (20) as follows, where we have substituted b_0 for the term in parentheses there:

$$\Delta M = M \cdot \left(b_0 \cdot \frac{\Delta E}{E} + \beta_{M,V} \cdot \beta_{V,PV} \cdot \frac{\Delta P_V}{P_V} \right) . \tag{33}$$

Observe that, since $\Delta E_3 = 0$, we cannot define a quantity b_3 comparable to b_1 and b_2. This is because a change in the vehicle price has no direct effect on fuel efficiency. This does not really matter for our purposes, as no attempts would be made to estimate the rebound effect only on the basis of variations in vehicle prices. However, the importance of controlling for vehicle prices in estimating the rebound effect, as noted before, remains. Furthermore, the full effect of a policy on VMT needs to include any affect arising through vehicle price, which can be expressed as a fraction of initial fuel consumption as follows:

$$B_3 = \frac{\Delta M}{M} = \left(b_0 \cdot \varepsilon_{E,PV} + \beta_{M,V} \cdot \beta_{V,PV} \right) \cdot \frac{\Delta P_V}{P_V} \tag{34}$$

where εE,PV is an elasticity expressing how choices of energy efficiency respond to vehicle price through changes in vehicle stocks and hence mileage.

---

[17] In as far as the vehicle price change follows from a regulatory change, the analysis pertains to the case where the regulatory change has no effect on the menu of fuel efficiencies (i.e. we exclude technology forcing standards), so that only a change in the sales mix, obtained through a change in vehicle prices, can guarantee compliance.

Each of sections 3.2.1, 3.2.2, and 3.2.3 are based on a change in a different exogenous variable. Furthermore, each of the VMT effects described in the resulting equations (20), (27), and (33) (each with the definition of ΔE given in that same subsection) is a first-order approximation valid for small changes. Therefore, following the rules of partial differentiation, the total change in VMT is just the sum of these individual changes. In this way, one can analyze policies that simultaneously change two or even all three of these variables, for example a regulatory policy that encourages more fuel-efficient vehicles and raises their price.

4.   Data Sets

In this section, we discuss the kinds of data sets that are available to analyze the rebound effect in California. We organize the discussion in the same way as the literature review: data sets suitable for aggregate time series, aggregate cross-sectional time series studies, and disaggregate cross-sections or panels.

4.1   Aggregate Time Series Data

The kinds of data used by Greene (1992), Jones (1993), and Schimek (1996) are from standard sources and so can be updated to the present time. Sources are described in Greene (1992, Appendix B) and Schimek (1996, p. 84). Data on VMT, fuel efficiency, and number of people with drivers' licenses are taken from the annual publication, *Highway Statistics*, issued by the Federal Highway Administration (FHWA). Gasoline prices for years starting 1977 are from *Monthly Energy Review* by the Energy Information Administration (EIA) of the US Department of Energy (DOE); for earlier years, Greene uses another EIA publication, *Annual Energy Review*, while Schimek uses *Platt's Oil Price Handbook and Oilmanac* and the American Petroleum Institute's *Petroleum Facts and Figures*. Aggregate economic activity (e.g. gross national product, national income) are easily taken from the national income accounts, maintained by the Bureau of Economic Analysis (part of the US Department of Commerce), and employment data are reported regularly by the Bureau of Labor Statistics (part of the US Department of Labor); the national statistics on economic activity and employment are reprinted in many sources, such as the *Statistical Abstract of the United States*.

The treatment of light trucks used as passenger vehicles causes some comparability problems between years before and after 1966. Starting in 1966, FHWA distinguished 2-axle, 4-tire trucks from other trucks; the former can be used as a proxy for light trucks used as passenger vehicles. Before that, the distinction was not made so researchers had to impute "light trucks" using assumptions along with data on all single-unit trucks. This affects the data for both VMT and fleet-average fuel efficiency.

Data on VMT are, of course, crucial to measurements of the rebound effect. For the most part these are estimated by FHWA from reports by each of the individual states, which in turn base their reports on traffic counts along samples of highway segments. Converting such counts to VMT involves many assumptions, and practice has been evolving over time. The procedures have become considerably more sophisticated since the beginning of the Highway Performance

Monitoring System (HPMS) in 1979; but still one cannot be certain that procedures are fully consistent from one year to the next. This is a limitation on the use of pure time series data, since there is no internal way to check for such problems without severely reducing the number of independent data points available for analysis of the effects of interest.[18]

4.2   Aggregate Cross-Sectional Time Series Data

The papers using aggregate cross-sectional time series data use similar variables and data sources to those for aggregate time series, but require that many variables be measured separately for each state (plus District of Columbia). In addition, these studies have collected data on other variables, which of course are then also available for pure time series studies.

Haughton and Sarkar (1996) provide specific data sources in their Appendix, covering years 1970-1991. Like the time-series studies, they use FHWA's *Highway Statistics* as a primary source because it provides state-specific data for most variables. They take national income and price indices from the *Statistical Abstract of the United States*, and vehicle registrations from *Ward's Automotive Yearbook*.

As alternate income measures, we consider disposable income (which is personal income less taxes) and gross state product, which is an aggregate measure of economic activity. Disposable income by state is available starting in 1969, and gross state product starting in 1977. When we use these latter two measures, we measure the ratio of that measure to personal income averaged over the ten earliest years we have the variable, and apply that ratio to personal income for the earlier years back to 1966.

Gasoline prices are also taken from the *Statistical Abstract of the United States*, which in turn gets them from different sources in different years: American Petroleum Institute before 1955, then *Platt's Oilgram Price* Service. These sources are nationwide only; but starting in 1970 they are available by state from the EIA on-line data series, *State Energy Data: Price and Expenditure* Data, Table 5.[19] A different series is available from EIA for 1960-1977, also by state.[20]

Another possible source for gasoline prices is the Consumer Price Index, compiled by the Bureau of Labor Statistics (BLS). These are not available by state but rather for 27 major metropolitan areas. Thus many but not all states would be partially accounted for, the more so the more urbanized the state. These data go back at least to 1976 (for 24 of the 27 metro areas); we are currently researching the status of this series for earlier years.

---

[18] For example, Greene estimates some models omitting data prior to 1966. For others, he uses all years but includes a dummy variable for years before 1966 in order to account for data differences.

[19] Haughton and Sarkar (1996) evidently obtained these state data from the EIA publication, *Petroleum Marketing Monthly*.

[20] Source: US Energy Information Agency (1979), Table B-1, "Gasoline Service Station Retail Prices".

Although existing studies have not done so, it would be possible to supplement this type of information with vehicle registration data compiled by a commercial source such as R.L. Polk & Co. Such data can be compiled in a number of different ways (at different costs), making it possible to estimate the number of new vehicles and perhaps the age of existing vehicles in any given year and state. The proportions of vehicles of various makes and models are also available. Of course such data will not reflect the nature of unregistered vehicles, which are numerous in California. Our judgment is that there would not be much advantage in adding data from this source unless one were going to estimate separate usage models for different types of vehicles.

There are several advantages of aggregate cross-sectional time series data sets. First, such data provide direct evidence of the aggregate effects which are of primary policy interest, without having to rely on samples of individuals and aggregation procedures. Second, they provide up to 51 times as many observations as a pure time series, allowing the analyst to take advantage of differences in conditions across states in order to help identify the subtle effects sometimes required; in particular, as noted earlier, such data appear capable of distinguishing between autocorrelation (correlation over time) and lagged effects (which are crucial to distinguishing short-run from long-run phenomena). A third advantage of a aggregate cross-sectional time series data set is that it includes California specifically, thereby permitting special attention to differences between California and other states that may appear from statistical analysis. This is especially important for our research mandate, which asks for California-specific estimates of the rebound effect. A method allowing systematic comparisons across states is preferable to one where results from different models need to be compared. Fourth, such data can be supplemented using other sources, if desired, in order to bring in additional variables that may help distinguish some of the channels of causality through which the rebound effect operates. Fifth, it is likely that such data sets can be extended forward in time as newer data are collected, making it possible to update any such study in future years.

There are also some disadvantages. One is possible inconsistency in data collection methods across states, especially for VMT. Second, some variables are available only at the national level. Third, the need to go back as far as possible in time, in order to provide the size data set needed to identify subtle effects, means that various definitions and sources change over time; this disadvantage is of course shared with pure time series. Fourth, the data require quite a bit of work to collect because they are from multiple sources and must be checked carefully for consistency.

Our overall assessment is that the disadvantages are not overly serious. Inconsistency across states and inability to measure some variables by state can be compensated for using standard statistical techniques for cross-sectional time series: namely, either *fixed effects* or *random effects* to represent unknown sources of error for each state. Similarly, inconsistency over time can be represented by dummy variables to represent possible effects of a change in measurement across a known time threshold. It is the size of this type of data set that makes these types of adjustments possible; because such a data set contains both extensive time-series and cross-sectional information, it is possible to reduce the degrees of freedom in either or both the time and spatial dimensions and still have enough degrees of freedom left to get good precision in statistical measurement.

4.3   Disaggregate Data

A number of micro data sets are available, either for a single year or for a small number of years. We review first nationwide data sets, then California-specific data sets.

4.3.1 National Disaggregate Data Sets

NPTS, NHTS, and CSHTS. The Nationwide Personal Transportation Survey (NPTS) is a sample of individuals or households that is widely used to track travel trends. FHWA has produced a long series of reports tabulating and describing the import of various statistics such as trip rates, commuting times, vehicles owned, vehicle-miles traveled, and how they vary across socioeconomic and demographic subgroups. A recent example of such a report is Hu and Young (1999), which covers the 1995 NPTS. In 1995, the sample size was 42,033 households with 95,360 separate people interviewed (Hu and Young 1999, p. 2). As described earlier, Pickrell and Schimek (1999) use the 1995 NPTS to estimate vehicle usage models, which among other things measure the rebound effect.

The NPTS was performed in 1969, 1977, 1983, 1990, and 1995. Starting in 2001, it was merged with another survey covering households (American Travel Survey, sponsored by the Bureau of Transportation Statistics) into the National Household Travel Survey (NHTS).

These surveys have the advantage of being large, regular, and well supported administratively. However, there have been regular problems of comparability across years in the variable of greatest interest here, namely personal vehicle-miles traveled. In 1990, the survey changed from personal to telephone interviews, which may have resulted in an under-representation of low-income households. Possibly as a result, the NPTS data show an increase in VMT of 50 percent from 1983 to 1990, much larger than 33 percent based on independently estimated FHWA data from *Highway Statistics* (Vincent, Keyes and Reed 1994, pp. 4-5). In 1995, three separate estimates of VMT became available, none precisely comparable to 1990; see Pickrell and Schimek (1999, pp. 2-4 and Appendix) for an analysis. Pickrell and Schimek provide adjustments to the 1990 data to make them more comparable to 1995, but this is at an aggregate level only. In addition to these problems, there has long been concern that survey respondents' recollection of VMT does not match independent checks using odometer readings.

The California subsample of the NHTS is being analyzed in a related research project at UC Irvine carried out by Thomas Golob. That research will attempt an analysis somewhat like that of Greene, Kahn, and Gibson (1999) described earlier, in which household fuel consumption is a function of the number and types of vehicles owned by that specific household as well as of demographic and socioeconomic characteristics and residential location.

There is a similar data set conducted by California known as the California Statewide Household Travel Survey (CSHTS), most recently carried out for 2000-2001. The survey design and an extensive presentation of descriptive results are described in NuStats (2002).

Both the 2001 NHTS and the 2000-2001 CSHTS seem to be good cross-sections for estimating models of individual household behavior. Such models can be the basis of simulation

models which trace household decisions through a number of dimensions including vehicle purchase, vehicle retirement, and individual vehicle usage. Projects of this type are quite ambitious in scope but offer great flexibility in analyzing different policy scenarios and tracing the effects of such scenarios on household vehicle fleet composition and usage.

Consumer Expenditure Survey (CES). This source was used by Goldberg (1996). The CES is conducted by the Bureau of Labor Statistics (BLS) as part of its work to construct the Consumer Price Index. It is performed quarterly, using about 4500 households each time. It is a continually refreshed panel, i.e. about three-fourths of each sample is reinterviewed as part of the next survey. Questions include sociodemographics, income, employment, and expenditures on various categories including new vehicle purchases. There is no direct measure of VMT; instead, amount of travel must be inferred from gasoline purchases and the fuel efficiency of the type of vehicle reported as being purchased; thus it is available only for households that purchase a vehicle sometime while they were in the sample.

Goldberg (1995, pp. 906-907) describes some other problems with the CES, which are relatively minor for us. It apparently underestimates new car purchase, perhaps because it excludes fleet sales.

The CES does not identify individuals by state of residence, making it impossible to derive a California-specific sample.

Residential Transportation and Energy Consumption Survey (RTECS). Since 1978, the EIA has collected a series of surveys, approximately every three years, on residential energy consumption, known as the Residential Energy Consumption Survey (RECS). Prior to 1983, this survey included a Household Transportation Panel with some information on vehicles and vehicle use. Starting in 1983, the much more complete RTECS was added, with data on number and types of motor vehicles owned, VMT on each vehicle, gasoline purchases, gasoline prices (either actual paid or, in later years, a matched consumer price index), household income, region of residence (the nine standard Census regions of the US), and degree of urbanization of residence (urban, suburban, or rural). Fuel efficiency of each vehicle is imputed either from gasoline purchases and VMT or from the vehicle type. The data are used and described more fully by Greene et al. (1999, pp. 10-11). The RTECS was discontinued after 1994. In 2001, EIA joined forces with the Department of Transportation to add energy-related questions to the NHTS.

The 1988 and 1991 surveys are used by Schmalensee and Stoker (1999) to estimate gasoline demand. In 1994 there were about 3,000 households sampled, owning about 5,500 vehicles.[21] In 1988 and 1991 there were about 2600-2700 households that reported having at

---

[21] This information is from the EIA web site, http://www.eia.doe.gov/emeu/rtecs/contents.html, accessed October 2003.

least one car and using it; 547 of these households were in the sample for both years (Schmalensee and Stoker 1999, p. 647).

The RTECS is the only survey data set for which VMT are obtained consistently by having respondents record odometer readings, a more accurate basis than asking respondents to remember. For this reasons it avoids the many problems with VMT data described earlier and also present in the California data described next.

The sample size, the repeated nature of the survey, the detailed information on vehicles, and the accurate VMT measurement make this data set an excellent one for estimating rebound effects. As described in Section 2, this has been carried out by Greene et al. (1999) with considerable success. Unfortunately the survey was discontinued after 1994, so it is not possible to update that study to more recent years. Nor is it possible to identify California residence, although one might try to infer it from other variables such as region, income and urbanization.

4.3.2 California Disaggregate Data Sets

Bureau of Automotive Repair. For some years, the Bureau of Automotive Repair has maintained data from its smog check program, which includes odometer readings on specific vehicles, identified by vehicle information number (VIN), along with test dates and owner zip code of residence. These data report on about 10 million tests per year. In earlier years, it was possible to match successive tests of the same vehicle in order to construct VMT between two test dates (usually about two years apart), but only with a very large amount of labor. From knowledge of the VIN, one can impute information about the vehicle type and age. Very recently the BAR has begun providing matched data for a subset of these vehicles, including vehicle type and VMT.

These data would need to be matched to other sources in order to estimate disaggregate models using socioeconomic, demographic, and price variables. Such matching is probably possible, but it is very labor intensive and would create considerable inaccuracy due to the need to use residential zip code as the only identifier of the individual. The data are subject to other severe limitations. Many cars change hands over two years, so the VMT recorded might not all be undertaken by the current owner. In addition vehicles enter and leave the data base through out-of-state migrations, vehicle retirements, failures to get the legally required smog check, erroneous data, and other sources, and these processes are not likely to be independent of the travel decisions we are interested in modeling. Correcting for such problems would involve even further assumptions and avenues for inaccuracy. Another UC Irvine research project, from the early 1990s, found that only 15 percent of vehicles in the BAR data at that time could be matched with a subsequent record for the same vehicle in order to create VMT; this would introduce considerable problems of selection bias. While not necessarily impossible, we judge this approach to require a major research effort and even so to entail some risk of bias in the results.

California Vehicle Survey (CVS). This survey, conducted in 2002, is intended to provide data for updating the CALCARS model used by the California Energy Commission (CEC). It consists of two parts: a revealed preference survey ("Recruit Survey") asking basic information including about the current household vehicle fleet and its usage, followed by a more extensive stated preference survey ("SP survey") in which respondents answer a series of questions about what vehicle they would choose in various hypothetical situations. Unlike most data sets, the CVS also includes a survey of owners of commercial vehicles.

The Recruit Survey concentrates on the current vehicle fleet and planned replacements. The usage questions include annual mileage for current vehicles, and so could be used as a source of estimates of the rebound effect. Extensive information is collected about likely retirements and replacements, making the survey especially well suited to models of vehicle purchase.

The SP Survey is almost entirely about vehicle purchase. It provides rich details about the characteristics of vehicles for prospective purchase, and so provides a good tool for measuring consumer response to changes in such characteristics, as well as to changes in price.

This kind of information is ideal for large-scale microsimulation models. The individualized information and the ability to ask about vehicles with traits not now on the market enable the analyst to consider numerous behavioral adaptations to regulations that cannot be observed in data on actual behavior. The obverse of this advantage is that such adaptations are unverified until there is enough actual experience to see whether people make the same choices they say they will.

4.4   Data Set Constructed for This Study

From our review of prior work and available data sets, we believe the most useful empirical approach for this project is to construct an aggregate cross-sectional time series data set of the type reviewed in Section 4.2. The advantages of an aggregate panel data set are set out there, as well as ways to overcome its disadvantages. Here, we elaborate on two of the points made there, with emphasis on the research goal of obtaining empirical estimates that are particularly relevant for California and for the time period in which CO_2 regulations might go into effect.

First, the cross-sectional time series approach offers a quite flexible way to investigate how variables that may differ between California and other parts of the US affect the rebound effect. This has not been done in prior studies, but is quite straightforward at least in the simpler models. Consider the partially reduced form VMT equation (11), in which VMT is explained as a function of per-mile fuel cost, among other things. This is the specification most closely resembling those commonly used in the literature. In such an equation, the rebound effect is computed from the coefficient of per-mile fuel cost P_M. But another variable could be included, in which P_M is multiplied by income, an urbanization measure, or other variables. In this way the rebound effect becomes not a constant, but a function of those variables. Of course, it remains to be seen whether any such effects are strong enough to be observed above the noise that is present in any statistical study. But a great advantage of a cross-sectional time series over a pure time

series (whether for US or California) is that its much greater size provides more data to find such effects above the noise.

This flexibility also helps to project rebound effects forward in time to a date when regulations might go into effect. For example, suppose the rebound effect depends on income or on degree of urbanization. If such dependence can be measured using the method of the previous paragraph, then projections of income and urbanization for future years can be applied to the estimated equation to see how large the rebound effect is likely to be under such projections.

Second, the cross-sectional time series approach allows one to single out California for its own special treatment. Of course this could be done by restricting the data set to California only, but that would bring in all the problems described for a pure time series. With a cross-sectional time series, California can be allowed to have some differences from the US without requiring that it be totally different. This can be done by including a dummy variable for California, both as a constant (to capture any California-specific traits leading to more or less VMT than its demographics would indicate) and, even better, interacted with selected other variables (to capture differences in how those variables affect VMT). The most important such interaction would be with fuel cost per mile, thereby allowing the rebound effect to differ by some constant amount in California compared to other states with the same demographics. Other interactions can be included as well, enabling the effects of other variables on VMT to differ between California and elsewhere.

If every possible such interaction were included, it would be equivalent to estimating a separate time series just on the California sample. This is an acceptable procedure logically, but unlikely to be the best one for reasons already given. Rather, it is better to use the full data set to help identify parameters that are not likely to be California-specific, like autocorrelation coefficients or effect of income, so that the more important variations can be more precisely estimated. This is what we can do by having a complete cross-sectional time series of many states instead of data just from California.

Third, the ability to update the data set by simply adding new information for later years is significant for meeting the goals of the Air Resources Board and Energy Commission. Because regulations will take effect several years in the future, there will no doubt be a desire to revisit questions about their effects as the implementation date draws closer. At that time several more years of data would be available, and it would be a relatively simple matter to add them to the data set and reestimate the same models, obtaining additional precision and confidence that the results are current.

A final advantage of an aggregate cross-sectional time series for this project is that it complements work being done at UC Davis under the sponsorship of the Air Resources Board. That work is developing a suite of microsimulation models which will permit extensive policy simulation, with detailed information about the nature of the responses. Microsimulation is the only way to answer many questions of interest in developing policy, and it also can estimate rebound effects by aggregating up from simulated individual responses. The current project will provide a check on such work, enabling such estimates to be compared with estimates based on actual behavior at an aggregate level. A check such as this is very useful because the kinds of assumptions needed for quantitative results are very different in the two methods, and by

examining the sources of any discrepancies it is possible to correct those assumptions and better understand the effect they have on results.

Our empirical work therefore uses cross-sectional time series data compiled by us from public sources, covering all 50 states of the US plus the District of Columbia, over the period 1966-2001. The specific variables are described in section 5.2. Data sources and some descriptive statistics are given in Appendix A.

5. Empirical Models on California-Relevant Data

5.1 System of Simultaneous Equations

In this study, we estimate the full structural model based on system (8), consisting of three simultaneous equations explaining VMT, vehicle stock, and fuel efficiency. We divide VMT and vehicle stock by number of adults, so in both cases we are explaining quantities per adult in the population. We also express all three dependent variables (and many independent variables as well) in natural logarithms because this seems a more plausible relationship and it is easy to interpret results as elasticities.

The equation system is generalized in two ways to handle the dynamic dimensions of observed statewide averages of these three dependent variables. First, we assume that the error terms in the empirical equations exhibit first-degree serial correlation; that is, they are autoregressive of order 1, known as AR(1). Second, we assume that each variable being explained exhibits inertia that is represented by including among the explanatory variables a single one-year lagged value of that dependent variable. The coefficient of this variable determines the difference between short-run and long-run effects of any of the independent variables.

Formally, then, the system is the following:

$$(vma)_t = \alpha^m (vma)_{t-1} + \alpha^{mv}(vehstock)_t + \beta_1^m (pm)_t + \beta_3^m X_t^m + u_t^m$$

$$(vehstock)_t = \alpha^v (vehstock)_{t-1} + \alpha^{vm}(vma)_t + \beta_1^v (pv)_t + \beta_2^v (pm)_t + \beta_3^v X_t^v + u_t^v \qquad (35)$$

$$(fint)_t = \alpha^f (fint)_{t-1} + \alpha^{fm}(vma)_t + \beta_1^f (pf)_t + \beta_2^f (cafe)_t + \beta_3^f X_t^f + u_t^f$$

with error terms following the rule

$$u_t^k = \rho^k u_{t-1}^k + \varepsilon_t^k, \text{ k=m,v,f.} \qquad (36)$$

Here, lower-case notation indicates that the variable is in logarithms. Thus vma is the natural logarithm of VMT per adult; vehstock is the log of number of vehicles per adult; and fint is the log of fuel intensity, defined as the reciprocal of fuel efficiency, or equivalently fint is the negative of the log of fuel efficiency. (Explaining equations in terms of fuel intensity rather than fuel efficiency is purely a matter of convenience, and makes no difference in the properties of the

estimators.) Variable pf is the log of fuel price; hence log fuel cost per mile, pm, is equal to pf+fint. The parameter β_1 is the coefficient of the log of a price (hence is a price elasticity), β_2 is the coefficient of a single additional variable (in log form); whereas β_3 is a vector of coefficients of the set of variables (including a constant) in the corresponding list X, which may be either in levels or logarithms. Subscript t designates a year, and u and ε are error terms assumed to have zero expected value. Furthermore ε is assumed to be "white noise," that is, independent and with the same distribution from each year to the next.

An example may clarify what all this means. The logarithm of vehicle usage per adult (vma) depends on its own previous value (representing inertia in usage decisions) and on the number of vehicles per adult (vehstock). It also depends on fuel cost per mile (pm), other variables (Xm), and unobserved influences (um). Two of these explanatory variables are endogenous : vehstock and pm, the latter because it is determined from fuel intensity fint (along with fuel price). Therefore these two variables require special treatment in the econometric procedure used to estimate the equation. The correlation in error terms, represented by (36), means that some of the unobserved factors influencing usage decisions in a given state will be similar from one year to the next: for example, laws governing driving by minors.

Because our variables are in logarithms, the coefficient of variable pm in the usage equation, β_1^m, is the same as εM,PM, which is by far the most important part of the equation (12b) defining the rebound effect. However, in addition to the other small terms in equation (12b), there are two other features of our empirical specification that modify the rebound effect. The first is that we include some variables in Xm that are interactions of pm with income or urbanization, so that the rebound effect varies with these measures. We do so in such a way that the coefficient of pm remains the same as εM,PM at the mean values of income and urbanization. In our tables of results, we show the result of calculating equation (12b) exactly, both at the mean values of variables inc in our sample and at the mean values for California in years 1997-2001. Using the notation of (35), (12b) takes the form:

$$-b^S = \varepsilon_{\tilde{M},PM} = \frac{\beta_1^m + \alpha^{mv}\beta_2^v}{1 - \alpha^{mv}\alpha^{vm}} \tag{37a}$$

where the symbol bS designates the short-run rebound effect.

The second feature modifying the rebound effect is the inclusion of lagged values. The coefficient on lagged vma (αm) in the usage equation indicates how much a change in one year will continue to cause changes in subsequent years, due to people's inability to make fast adjustments in lifestyle. For example, a change Δpm in fuel cost per mile in year 1990 would cause a change in vma of $\beta_1^m \cdot \Delta pm$ in 1990, $\alpha^m \beta_1^m \cdot \Delta pm$ in 1991, $(\alpha^m)^2 \beta_1^m \cdot \Delta pm$ in 1992, $(\alpha^m)^3 \beta_1^m \cdot \Delta pm$ in 1993, and so forth, converging ultimately to a change $[\beta_1^m/(1-\alpha^m)] \cdot \Delta pm$ after a long time.[22] Ignoring the small indirect effects via the equation for vehicle stock, then, we can identify β_1^m as the short-run rebound effect and $\beta_1^m/(1-\alpha^m)$ as the long-run rebound effect,

---

[22] This is due to the mathematical fact that an infinite sum $(1+\alpha+\alpha^2+\alpha^3+...)$ equals $1/(1-\alpha)$.

in both cases at the mean values of inc and Urban in the data set. The precise equation for the long-run rebound effect is of a form similar to (37a) but with different multipliers on the β coefficients:

$$-b^L = \varepsilon^L_{\tilde{M},PM} = \frac{\beta_1^m \cdot (1-\alpha^v) + \alpha^{mv}\beta_2^v}{(1-\alpha^m)(1-\alpha^v) - \alpha^{mv}\alpha^{vm}} \, . \tag{37b}$$

(We demonstrate this in Section 6 by considering a dynamic system.)

The same considerations apply to other elasticities. Using comparable computations, one can show that the short- and long-run elasticities of vehicle usage with respect to new-car price are:

$$\varepsilon^S_{\tilde{M},PV} = \frac{\alpha^{mv}\beta_1^v}{1-\alpha^{mv}\alpha^{vm}} \, ; \qquad \varepsilon^L_{\tilde{M},PV} = \frac{\alpha^{mv}\beta_1^v}{(1-\alpha^m)(1-\alpha^v) - \alpha^{mv}\alpha^{vm}} \tag{38}$$

and the short- and long-run elasticities of fuel intensity with respect to fuel price are:

$$-\varepsilon^S_{\tilde{E},PF} = \frac{\beta_1^f + \alpha^{fm}\beta_1^m}{1-\alpha^{fm}\beta_1^m} \, ; \qquad -\varepsilon^L_{\tilde{E},PF} = \frac{\beta_1^f \cdot (1-\alpha^m) + \alpha^{fm}\beta_1^m}{(1-\alpha^f)(1-\alpha^m) - \alpha^{fm}\beta_1^m} \, . \tag{39}$$

Equations (39) are approximations in which just the two-way causation between fint and vma is accounted for, rather than also including the even more indirect effect of pf on fint via the effect of vehicle stock on vehicle usage combined with the effect of vehicle usage on fuel intensity (this effect which will be especially small because it involves the triple product $\beta_2^v \alpha^{mv} \alpha^{fm}$).

Similarly, we can approximate the income elasticity of vehicle usage by ignoring the short-run income-elasticity of vehicle efficiency, which turns out to be extremely small in our estimates. Then we need only account for the first two of equations (35). The elasticity is:

$$\varepsilon^S_{\tilde{M},Y} = \frac{\beta_{3Y}^m + \alpha^{mv}\beta_{3Y}^v}{1-\alpha^{mv}\alpha^{vm}} \, ; \qquad \varepsilon^L_{\tilde{M},Y} = \frac{\beta_{3Y}^m \cdot (1-\alpha^v) + \alpha^{mv}\beta_{3Y}^v}{(1-\alpha^m)(1-\alpha^v) - \alpha^{mv}\alpha^{vm}} \tag{40}$$

where β_{3Y}^m and β_{3Y}^v are the coefficients of log income in the usage and vehicle stock equations.

5.2   Specification of the Equations

In this subsection, we describe the variables we use and their rationale, as well as issues arising in a cross-sectional time series data set. We pay special attention to a variable describing CAFE regulation, which we identified in the literature review as lacking theoretical foundation and possibly being responsible for disparate results. In each case we give first the corresponding notation of the theory section, and end with the variable name used in our base regressions (variants are described along with results). Variables starting with lower case letters are logarithms of the variable described. Data sources are given in the Appendix.

A few studies have tried to estimate separate effects for increases and decreases in fuel price per mile. We do not find these attempts very convincing because it has been quite difficult to approximate an appropriate dynamic model within what is basically a static framework. Furthermore, any asymmetric responses would be short-run, and would probably not affect the longer-run properties of the system because new desired characteristics can gradually be incorporated into the vehicle fleet as it turns over, which happens within about one decade.

5.2.1 Dependent Variables

M: Vehicle miles traveled (VMT) divided by adult population, by state and year (logarithm: vma, for "vehicle-miles per adult").

V: Vehicle stock divided by adult population (logarithm: vehstock).

1/E: Fuel intensity, calculated as highway use of gasoline divided by VMT (logarithm: fint).

5.2.2 Independent Variables other than R_E

PM: This variable is simply the ratio of the real price of fuel to E. Its logarithm is denoted pm $\equiv$ ln(PF)–ln(E) $\equiv$ pf+fint. For convenience in interpreting interaction variables involving pm, we have normalized pm by subtracting its mean (6.8140) for the sample. (More precisely, we normalize its components, pf and fint.)

XM: This set of variables includes the following:

Real personal income per capita at 1987 prices,[23] in log form and normalized by subtracting the sample mean of (inc); number of adults divided by public road mileage (logarithm: adrm) as a rough measure of urbanization or potential congestion that is less likely to be endogenous than actual congestion; ratio of total population to adults (logarithm: popratio) as a measure of family size; fraction of state's population living in metropolitan statistical areas (Urban), normalized by subtracting its mean in the sample; fraction of the state's population living in metropolitan statistical areas with a heavy-rail transit system (Railpop); a dummy variable to represent gasoline supply disruptions in 1974 and 1979 (D7479); and a time trend measured in years since 1966 (Trend). We hope the time trend captures some of the changes in technology and consumer preferences that we are unable to specify quantitatively.[24] We

---

[23] Real income and costs are stated in 1987 prices because 1987 is one of the base years used for stating the consumer price index, and it is near the middle of our time period so that using it as a base year minimizes errors from changing quantity weights. To restate these values at 2002 prices, divide by 1.584 (since the 1987-based consumer price index was 158.4 in year 2002).

[24] Instead of the Trend, we have experimented with three technology variables: vehicle volume (*Vol*), engine horsepower (*Hp*), and top speed (*Speed*), each in the form a fractional change in that measure since 1975, the earliest year for which we have the measure, and zero prior to 1975. For years 1966-1974, we include a trend variable *Techtrend* equal to min{(year-1975), 0} in order to capture the effects of any earlier changes (assumed linear) in these variables. This experiment has not yet proven fruitful.

also interact pm with inc (or with dispinc or gsp) and with Urban in order to test the hypothesis that the cost elasticity declines as time costs become a more prominent part of the cost of driving, which could happen either because those costs are valued more (as incomes rise) or because they are larger (because of urban congestion). We also tried other variables, including the fraction of a state's population representing growth since 1950, and an interaction between pm and a California dummy variable; neither of these variables showed any explanatory power.

All income and price variables are deflated by a national consumer price index, rather than a state-specific index, because the only local variation in consumer prices available is for metropolitan areas; applying such an index to an entire state would create potential biases because some states have a much greater fraction of their VMT in metropolitan areas than others. Furthermore, as noted in the introduction, the critical role of income in our study is expected to be through the value of time, which raises the non-fuel cost of travel and therefore reduces the proportion of cost that is accounted for by fuel. Value of time will be high in states with high nominal income, regardless of whether they also have high cost of living.

As alternative measures of income, we considered disposable income (personal income after taxes) and gross state product (a measure of value of production, which unlike personal income includes the business sector). They are approximated for earliest years in the sample as described in Section 4.2. These variables, like personal income, are put in log form and then normalized by subtracting the corresponding sample mean. They are named dispinc and gsp; like inc, each is entered in the equation both by itself and interacted with pm. In each variant where inc is replaced by one of these alternate measures, the replacement is made throughout the system, i.e. in the set of variables XV and in the creation of variable cafe (both described below).

We have not specified a variable for operating costs other than fuel because they would vary little across states and only very slowly over time. Whatever role might be played by such costs is likely to be captured by the time trend variable.

PV: Index of real new vehicle prices (1987=100) (logarithm: pv).

XV: We include inc, adrm, and Trend, already defined in XM. In addition there are two other variables: the national interest rate for auto loans (logarithm: interest); and the ratio of licensed drivers to adults (logarithm: licad).

PF: Price of gasoline, real at 1987 prices (cents per gallon). Its logarithm (pf) is normalized by subtracting the mean in the sample.

RE: This variable is described in the next subsection.

XE: These variables include six variables in XM, namely inc, adrm, popratio, Urban, Railpop, and D7479. Instead of using a single linear time trend, we allow for the possibility of three distinct trends in fuel efficiency: one before the OPEC embargo (1966-1973), another between the embargo and the Iranian revolution (1974-1979), and a third after the Iranian revolution in 1979. The rationale is that these events changed people's perception of long-

term prospects for oil supplies and therefore affected research and development efforts related to fuel efficiency. On the assumption that changes in technology cannot happen immediately, these variables are specified in such a way that there is a break in the slope of the trend line but not a sudden "jump" from one regime to another. Specifically, we define Trend1 = Min (Trend, 7); Trend2 = Min (Trend, 11) – Trend1; and Trend3 = Trend – Trend1 – Trend2.

5.2.3 Variable to Measure CAFE Regulation

We define a variable intended to measure the tightness of CAFE regulation starting in 1978: specifically, the difference between the mandated efficiency of new passenger vehicles and the efficiency that would be chosen in the absence of regulation. This difference is truncated at zero, that is, the variable is zero when CAFE is not binding or when it is not in effect. We interpret it as a variable that influences the efficiency of new passenger vehicles, because the inclusion of a lagged dependent variable in the fuel-intensity equation already captures the inertia due to slow turnover of the vehicle fleet.

The calculation proceeds in four steps. (1) We estimate a reduced-form equation explaining fuel intensity from 1966-1977. (2) This equation is interpreted as a partial adjustment model, so that the coefficient γ of lagged fuel intensity enables us to form a predicted desired fuel intensity for each state in each year (from which actual fuel intensity is obtained by moving a fraction γ of the way from last year's value to the desired value, plus the random error term). This prediction can be done for all years in the sample by applying the values of the independent variables for those years. (3) For a given year, we average desired fuel intensity (weighted by vehicle-miles traveled) across states to get a national desired average fuel intensity. (4) We measure the strength of CAFE regulation by whether and how far the minimum mandated efficiency (corrected for the difference between testing equipment and real-world driving) exceeds the reciprocal of the national desired average fuel intensity. Specifically, after taking logarithms of both the mandated and desired fuel efficiency, the variable cafe is set equal to their difference if it is positive, or to zero if it is not. See Appendix A for further details.

Implicit in this definition is a view of the CAFE regulations as exerting a force on every state toward greater fuel efficiency of its fleet, even if that particular state has a desired fuel efficiency that meets the CAFE standard. The reason is that the standard applies to the nationwide fleet average for each manufacturer, and the manufacturer therefore has an incentive to use pricing or other means to improve fuel efficiency everywhere, not just where it is low.

5.2.4 Error Structure

Our data set is a cross-sectional time series, with each state observed 36 times. We follow conventional practice by allowing for the possibility that the error terms uit are not independent. We considered both fixed effects and random effects models, but found that a standard Hausman test easily rejects the latter in favor of the former. Random effects models are also prone to the omission of factors that influence our dependent variables but are correlated with variables we

include. Fixed effects models eschew using this cross-sectional information in favor of time series information, which in our case is abundant both because our time series is long (36 years) and because we have 51 different time series (although many of them are very similar to each other). Therefore we show only the fixed effects models in our results.

5.2.5 Autoregression and Autocorrelation

As noted earlier, consistent estimates of variables in a long cross-sectional time series data set may depend critically on correctly incorporating autoregression (dependence of a variable on its own lagged value) and autocorrelation (in which the error term is a moving average of independent error terms). These factors are especially important in determining the short-run and long-run elasticities, because they differ precisely when there is an autoregressive structure and the measurement of this autoregressive structure is sensitive to whether or not autocorrelation is controlled for. Indeed, one motivation for this study is that previous work has not been able to disentangle autoregression from autocorrelation with great confidence, and we believe our longer time series may make this possible. As already noted, this is important for determining short- and long-run elasticities, because they differ insofar as the lagged dependent variable has an important influence on current (short-run) behavior: specifically, the short-run response to an exogenous change will be conditioned on the most recent lagged values, while the long-run response is one in which those lagged values gradually change.

We have done some experimentation and determined to our satisfaction that a single lag is adequate to capture autoregression and that a first-order process is adequate to describe autocorrelation. Fortunately, our data are sufficiently abundant that both seem to be estimated with adequate precision. Furthermore, including both does not seem to detract from the precision of the other estimates. Therefore we have some confidence that the resulting estimate of the coefficient of the lagged endogenous variable is accurate and gives a valid indication of the extent of long-run effects.

We use the autocorrelation feature in the computer package Eviews 5, which estimates a model with first-order autocorrelation, as in (35)-(36), by transforming it to a nonlinear model with no autocorrelation but additional lags. This is done by writing the k-th equation lagged one year, multiplying it by ρ^k, then subtracting it from the unlagged equation and solving for the current dependent variable. The result is an equation whose right-hand side contains two lags (t-1 and t-2) of the dependent variable and one lag (t-1) of all the other variables. The coefficients of these variables are mixtures of the unknown parameters in (35) and (36), allowing all those parameters to be estimated using nonlinear least squares.

There is a simpler two-step procedure in which first the parameters of (35) are estimated using ordinary least squares and ρ^k is estimated from (36); that value of ρ^k is then used to transform (35) into a form that provides a better estimate. However, this procedure is known to be statistically biased when the model contains a lagged dependent variable, as ours does (Davidson and MacKinnon, p. 336).

Equations exhibiting values of ρ^k close to one are likely to be mis-specified, perhaps missing some of the most important explanatory variables (Davidson and MacKinnon, p. 335).

Fortunately we find that all our estimates of autocorrelation parameters are less than 0.25 in absolute value, indicating that the variables we measure are explaining the lion's share of effects that vary over time.

5.2.6 Estimation of the Simultaneous Equations

Two procedures are available for estimating systems like (35) containing several endogenous variables. The first, two-stage least squares (2SLS), first estimates a reduced form of the system somewhat analogous to (10);[25] it then estimates each equation while replacing the endogenous variables on its right-hand side by their predicted values from the first stage.

The other procedure, three-stage least squares (3SLS), goes a step further and estimates correlations in the error terms (36) among equations. This is likely in our system because, for example, unobserved factors like settlement patterns might influence both usage and vehicle stock. It then re-estimates the entire system taking these correlations into account. The advantage of 3SLS is that is makes more efficient use of the data, by taking advantage of the information in the correlations among the endogenous variables, and therefore permits more precise measurement of parameters. The disadvantage is that if there are errors in the specification of one equation (for example an omitted variable or a variable entered as a logarithm when it should be entered as a polynomial), then this error affects the other equations more directly than with 2SLS. We do not find any obvious problems with either of these procedures, and we show results with both; we consider the 3SLS results our best estimates and use them in Section 6.

5.3   Results

5.3.1 Structural Equations

The results of estimating the structural system are presented in Tables 1-3. Each table shows three different estimation methods. It is encouraging that there is little difference between three-stage least squares (3SLS) and two-stage least squares (2SLS); the former provides slightly better precision of estimates, as it theoretically should, and there are no signs of problems that might arise from mis-specification. We therefore accept the 3SLS results as our best estimates. The ordinary least squares (OLS) results are shown for comparison.

---

[25] More precisely, each equation contains as variables only the exogenous contemporary variables, but for technical reasons it must also contain one lagged value of all the exogenous variables and two lagged values of all three endogenous variables. See Fair (1984, ch. 6) or Davidson and MacKinnon (1993, ch. 10) for an explanation.

Table 1. Usage Equation

Variable	Estimated Using Three-Stage Least Squares		Estimated Using Two-Stage Least Squares		Estimated Using Ordinary Least Squares	
	Coefficient	Stndrd. Error	Coefficient	Stndrd. Error	Coefficient	Stndrd. Error
vma(t-1)	**0.7786**	**0.0133**	**0.7785**	**0.0149**	**0.7434**	**0.0153**
vehstock	**0.0538**	**0.0114**	**0.0332**	**0.0152**	**0.0556**	**0.0127**
pm	**-0.0521**	**0.0046**	**-0.0528**	**0.0048**	**-0.0860**	**0.0048**
pm*(inc)	**0.0837**	**0.0166**	**0.0920**	**0.0218**	**0.0682**	**0.0180**
pm*(Urban)	0.0050	0.0119	0.0103	0.0159	0.0135	0.0140
inc	**0.0933**	**0.0146**	**0.0964**	**0.0157**	**0.0955**	**0.0157**
adrm	**-0.0162**	**0.0063**	**-0.0194**	**0.0068**	**-0.0166**	**0.0068**
popratio	**0.1476**	**0.0363**	**0.1381**	**0.0405**	0.0460	0.0385
Urban	**-0.0447**	**0.0207**	**-0.0529**	**0.0223**	-0.0353	0.0224
Railpop	0.0021	0.0082	0.0030	0.0087	0.0009	0.0088
D7479	**-0.0437**	**0.0035**	**-0.0436**	**0.0036**	**-0.0360**	**0.0036**
Trend	0.0003	0.0003	0.0005	0.0004	-0.0006	0.0004
constant	**2.1014**	**0.1274**	**2.1142**	**0.1408**	**2.4991**	**0.1466**
rho	**-0.0954**	**0.0235**	**-0.0643**	**0.0288**	-0.0239	0.0294
No. observations	1,785		1,785		1,785	
Adjusted R-squared	0.9805		0.9806		0.9812	
S.E. of regression	0.0318		0.0317		0.0313	
Durbin-Watson stat	1.9196		1.9860		1.9870	
Sum squared resid	1.6861		1.6817		1.6333	

Notes: Bold or italic type indicates the coefficient is statistically significant at the 5% or 10% level, respectively.
Estimates of fixed effects coefficients (one for each state except Wyoming) are not shown.
Variables *inc*, *Urban*, and the constituent variables in *pm* are normalized by subtracting their mean value in the sample, both in the variable itself and in any interactions it takes. As a result, the coefficient of any variable in its uninteracted form gives the effect of that variable on *vma* at the mean values of the other variables.

The equation explaining usage (Table 1) is in many ways the most satisfactory. This equation tells how much driving is done by the average adult, holding the vehicle stock constant. Many coefficients are identified with good precision and demonstrate a strong and plausible effect. The direct short-run and long-run income elasticities are 0.093 and 0.421, respectively, at the average values for fuel cost per mile and urbanization; this long-run value is be compared to estimates of 0.2 to 0.5 in four studies compiled by Schimek (1996, Table 5). Each adult tends to travel more if there is a larger road stock available (as indicated by the negative coefficient on adrm), and if the average adult is responsible for more total people (popratio). While adrm may capture the effects of congestion, our measure of urbanization does not seem to have much effect, although it is in the expected direction. The proportion of population with rail transit available has no discernable effect, probably because it is too crude a measure of what transit options are really available. The two years 1974 and 1979 exhibited a lower usage, by about 4.4 percent, other things equal.

To help visualize the way income affects usage in the model of Table 1, we show in Figure 2 how rising income shifts the demand curve for vehicle-miles (conditional on vehicle stock) outward, while at the same time rotating it. The outward shift represents the income-elasticity of demand; the rotation represents the change in the rebound effect.

Figure 2. Effect of Income on Demand Curves for Vehicle-Miles
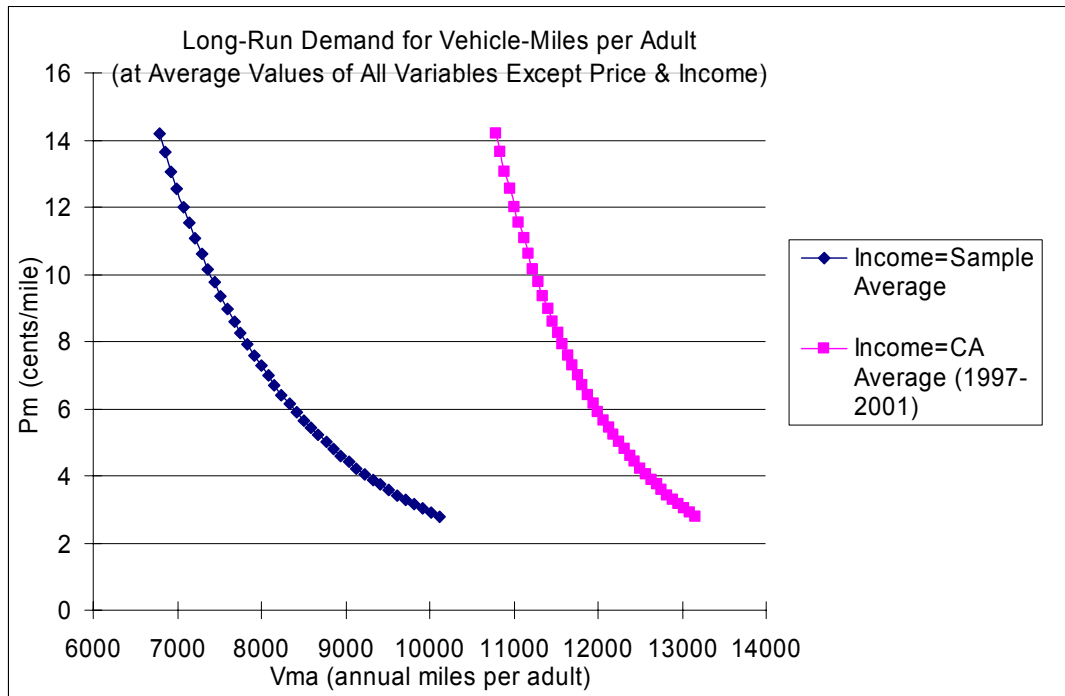
The vehicle stock equation (Table 2) is less satisfactory for purposes of tracking price effects because neither the price of a new car nor the cost of driving a mile have a significant effect on the vehicle stock. Income does have a significant effect (5% significance in a one-tail test), with rather low direct short- and long-run income elasticities of 0.028 and 0.182; and so does road provision (adrm) and a high proportion of adults having drivers' licenses (licad). It seems that the vehicle stock is better explained by basic characteristics of the population of potential car-owners and of the road infrastructure than by price variation. Of course, stronger variation in car prices than what is observed in our data may still significantly affect car ownership decisions. As expected, there is strong inertia in expanding or contracting the vehicle stock, as indicated by the coefficient of about 0.85 on the lagged value of vehicle stock. This means that any short-run effect, for example from an increase in income, will be magnified by a factor of $1/(1-0.85) = 6.7$ in the long run.

The equation for fuel intensity (Table 3) plausibly shows a substantial effect of fuel price, in the expected direction. It also suggests that CAFE regulation had a substantial effect of enhancing the fuel efficiency of vehicles. Urbanization appears to reduce fuel intensity, perhaps due to a preference for small cars in areas with tight street and parking space. The time trends show a break following 1979 toward more fuel-efficient cars if other things had remained equal (which of course they did not). Surprisingly, the period between 1974 and 1979 showed the opposite trend. All of these trends are less than one percent per year, so probably not too much consequence should be attributed to them.

Variable	Estimated Using Three-Stage Least Squares		Estimated Using Two-Stage Least Squares		Estimated Using Ordinary Least Squares	
	Coefficient	Stndrd. Error	Coefficient	Stndrd. Error	Coefficient	Stndrd. Error
vehstock(t-1)	**0.8477**	**0.0149**	**0.8466**	**0.0153**	**0.8425**	**0.0153**
vma	0.0155	0.0161	0.0175	0.0164	0.0341	0.0147
pv	-0.0437	0.0384	-0.0505	0.0393	-0.0430	0.0391
pm	-0.0084	0.0064	-0.0076	0.0066	-0.0016	0.0065
inc	*0.0277*	*0.0154*	*0.0264*	*0.0157*	0.0220	0.0155
adrm	**-0.0239**	**0.0069**	**-0.0234**	**0.0070**	**-0.0224**	**0.0070**
Trend	-0.0006	0.0008	-0.0008	0.0008	-0.0006	0.0008
interest	-0.0017	0.0070	-0.0038	0.0072	-0.0049	0.0072
licad	**0.0375**	**0.0153**	**0.0417**	**0.0157**	**0.0412**	**0.0157**
constant	-0.0268	0.1579	-0.0394	0.1613	-0.2014	0.1452
rho	**-0.1538**	**0.0281**	**-0.1504**	**0.0288**	**-0.1472**	**0.0289**
No. observations	1,785		1,785		1,785	
Adjusted R-squared	0.9638		0.9638		0.9638	
S.E. of regression	0.0365		0.0364		0.0364	
Durbin-Watson stat	1.9508		1.9546		1.9537	
Sum squared resid	2.2234		2.2227		2.2207	

Table 2. Vehicle Stock Equation

Notes: Bold or italic type indicates the coefficient is statistically significant at the 5% or 10% level, respectively. Estimates of fixed effects coefficients (one for each state except Wyoming) are not shown.

Study to Evaluate Effects of Reduced Greenhouse Emissions on Travel

	Table 3. Fuel Intensity Equation					
	Estimated Using Three-Stage Least Squares		Estimated Using Two-Stage Least Squares		Estimated Using Ordinary Least Squares	
Variable	Coefficient	Stndrd. Error	Coefficient	Stndrd. Error	Coefficient	Stndrd. Error
fint(t-1)	**0.7901**	**0.0187**	**0.8046**	**0.0230**	**-0.0758**	**0.0197**
vma	**-0.0635**	**0.0239**	**-0.0624**	**0.0270**	**-0.8426**	**0.0207**
pf	**-0.0549**	**0.0068**	**-0.0423**	**0.0072**	**-0.0747**	**0.0078**
cafe	**-0.1021**	**0.0118**	**-0.0766**	**0.0146**	**-0.2860**	**0.0210**
inc	0.0089	0.0183	0.0223	0.0190	**0.4005**	**0.0281**
adrm	-0.0093	0.0077	-0.0092	0.0079	**-0.0481**	**0.0183**
popratio	**0.1289**	**0.0498**	**0.1163**	**0.0554**	**0.7457**	**0.0857**
Urban	**-0.1521**	**0.0533**	*-0.1200*	*0.0653*	-0.2105	0.2268
Railpop	-0.0134	0.0099	-0.0129	0.0100	**-0.0452**	**0.0206**
D7479	**-0.0090**	**0.0045**	-0.0070	0.0047	**-0.0164**	**0.0025**
Trend1	0.0006	0.0011	0.0004	0.0014	0.0066	0.0050
Trend2	**0.0032**	**0.0013**	-0.0003	0.0014	-0.0010	0.0025
Trend3	**-0.0039**	**0.0004**	**-0.0034**	**0.0004**	**-0.0137**	**0.0008**
constant	-0.0583	0.2058	-0.0064	0.2279	**4.9574**	**0.2425**
rho	**-0.1218**	**0.0240**	**-0.1339**	**0.0289**	**0.7886**	**0.0168**
No. observations	1,785		1,785		1,785	
Adjusted R-squared	0.9610		0.9613		0.9793	
S.E. of regression	0.0394		0.0393		0.0288	
Durbin-Watson stat	1.9390		1.9609		2.2598	
Sum squared resid	2.5965		2.5768		1.3825	

Notes: Bold or italic type indicates the coefficient is statistically significant at the 5% or 10% level, respectively.
Estimates of fixed effects coefficients (one for each state except Wyoming) are not shown.

OLS is a particularly bad estimator of the fuel intensity equation (Table 3). It greatly underestimates the lag coefficient and instead attributes observed serial correlation in fuel intensity to a very strong autocorrelation pattern—so strong that if it were true, it would indicate serious omissions from the explanatory variables. Fortunately the 3SLS and 2SLS estimators show that in fact autocorrelation is modest, and the inertia in fuel intensity (lag coefficient 0.79) is nearly as large as that in the vehicle stock.

5.3.2 Rebound Effects and Other Elasticities

Table 4 computes the rebound effects (stated as the negative of the cost-per-mile elasticity of driving), as well as other elasticities implied by the structural models. As noted earlier, the interactions through the simultaneous equations modify only slightly the numbers that can be read directly from the coefficients. In particular, the average cost-per-mile elasticity in the sample is -0.0526, which is nearly identical to the coefficient of pm in Table 1. Our best estimate of the long-run elasticity in the average state over the time period of our sample is -0.2553. Thus the average rebound effect in this sample is estimated to be approximately 5.3 percent in the short run and 26 percent in the long run.

41

Table 4. Rebound Effect and Other Price Elasticities						
	Estimated Using Three-Stage Least Squares		Estimated Using Two-Stage Least Squares		Estimated Using Ordinary Least Squares	
	Short Run	Long Run	Short Run	Long Run	Short Run	Long Run
Elasticity of VMT with respect to fuel cost per mile: (a)						
At sample average	-0.0526	-0.2553	-0.0531	-0.2500	-0.0862	-0.3537
Calif. 1997-2001	-0.0220	-0.1133	-0.0181	-0.0896	-0.0587	-0.2415
Elasticity of VMT with respect to new veh price:	-0.0024	-0.0716	-0.0017	-0.0502	-0.0024	-0.0621
Elasticity of VMT with respect to income:	0.0948	0.4773	0.0973	0.4687	0.0969	0.4220
Elasticity of fuel intensity with respect to fuel price:						
At sample average	-0.0518	-0.2049	-0.0391	-0.1517	-0.0025	0.2614
Calif. 1997-2001	-0.0536	-0.2392	-0.0412	-0.1955	-0.0267	0.1329
Elasticity of fuel consumption with respect to fuel price:						
At sample average	-0.1044	-0.4602	-0.0922	-0.4017	-0.0887	-0.0922
Calif. 1997-2001	-0.0756	-0.3526	-0.0593	-0.2852	-0.0855	-0.1087

Note: (a) The rebound effect is just the negative of this number (multiplied by 100 if expressed as a percent)

Our estimate of the short-run rebound effect without correcting for endogeneity is quite close to the consensus of the literature, whereas with the correction it is somewhat lower than this consensus. We note that Greene et al. (1999), one of the very few previous studies that takes endogeneity of fuel efficiency into account, obtains a long-run rebound effect of 23 percent, very close to ours. (They do not estimate a short-run rebound effect.)

What is novel in our study is that the model for vehicle usage is able to discern an additional influence of real income on the rebound effect. The coefficient on pm interacted with logarithm of income (both normalized by subtracting the mean value over the entire sample) shows that each increase in the logarithm of income by 0.1 (roughly a ten percent increase in income) reduces the magnitude of the short-run rebound effect by 0.1 x 0.0837 = 0.00837, or just under one percentage point. This appears to confirm the theoretical expectation that higher incomes make people less sensitive to fuel costs. (Note that this calculation does not involve the coefficient of income itself, which is simply the income elasticity at average values.)

We also allowed the rebound effect to differ by degree of urbanization, but that estimated effect is essentially zero. In other estimates not shown, we included a variable allowing the rebound effect to differ in California from other states even aside from the influence of income and urbanization, but this variable was very small and statistically insignificant. We conclude, then, that income is the primary source of any difference between California and other states in the size of the rebound effect.

To get an idea of the implications of income for the rebound effect, we compute the elasticity of usage with respect to cost per mile for values of income and urbanization equal to those measured for California over the most recent five-year period covered in our data set, namely 1997-2001. These results are also shown in Table 4. Again using the 3SLS results, the short-run rebound effect is reduced to 2.2 percent and the long-run effect to 11.3 percent. About half of the difference between these results and those at the sample average is due to the difference between California and other states, and about half to the higher incomes prevailing in 1997-2001 than over the entire period 1966-2001. California is not an outlier in per capita income: in 1997-2001 it ranked 11th among the 51 states (including Washington, DC), with the highest-ranking state, Connecticut, exceeding California's per capita income by nearly 30 percent.

The bottom two panels in Table 4 provide information about how fuel prices affect fuel intensity and overall fuel consumption. The former effect is estimated with great precision, as seen in the small standard error on the coefficient of pf in Table 3. It implies that a ten percent increase in fuel price causes consumers to choose cars with 0.55 percent greater fuel efficiency in the same year, and $0.55/(1-0.79) = 2.6$ percent greater over a long period of time if the rise in fuel price were to persist. Adding the elasticity due to vehicle-miles traveled gives the total elasticity of fuel consumption, shown in the last panel of Table 4. This estimate of long-run price-elasticity of fuel consumption is -0.46, very close to the middle of recent studies. In fact, our estimates of both this long-run overall price-elasticity of fuel demand and the proportion of it due to changes in usage ($0.2553/0.4602 = 55$ percent) are very much in line with the literature; see the review by Parry and Small (2002), who choose as the best consensus an elasticity equal to -0.55, with 40 percent of it caused by mileage changes.

Because income plays a crucial role, we investigated two other measures of income to see if they changed the strong influence that we find for income on the rebound effect. The first is to substitute disposable income, which excludes taxes, for personal income. These results are barely distinguishable from those using personal income, so we do not present them here. The second is to use gross state product instead (GSP) of income. On the surface, this might seem preferable because some travel is business travel, and GSP includes business income. However, most travel is personal, not business-related, and would be influenced by the income accruing to residents rather than businesses. Not only that, GSP is a poor measure of the value of time of people doing the traveling because it includes commuters from out of state: indeed, the gross state product per capita of Washington, D.C., averages 2.8 times that of its two neighboring states. Many people living in Maryland or Virginia travel to jobs in Washington by public transit, so most of their automobile travel is likely to be in their home state, whereas their contribution to GSP is in Washington, D.C. More generally, the many people who commute across state lines undertake most of their personal travel and some of their commuting travel within their state of residence, so in order to capture the influence of income on this travel we would want the income to be measured in their state of residence. In addition, figures on GSP are not available for the first 11 years of our sample. Nevertheless, the results using GSP fit about as well as those using personal income, and they are presented as a comparison in Tables 5 and 6. These results show the rebound to be declining in GSP, but less so than with personal income. (Some of this difference is compensated by the fact that urbanization plays a stronger, though still small, role in this version of the model.) We do not think the results using GSP are a good basis for future projections so do not use them further.

Table 5. Comparison of Usage Equations

Variable	Estimated Using Personal Income		Estimated Using Gross State Product		Estimated Using Time Trend	
	Coefficient	Stndrd. Error	Coefficient	Stndrd. Error	Coefficient	Stndrd. Error
vma(t-1)	**0.7786**	**0.0133**	**0.7876**	**0.0153**	**0.7898**	**0.0137**
vehstock	**0.0538**	**0.0114**	**0.0496**	**0.0127**	**0.0464**	**0.0115**
pm	**-0.0521**	**0.0046**	**-0.0544**	**0.0048**	**-0.0598**	**0.0108**
pm*(inc, gsp, or trend)	**0.0837**	**0.0166**	**0.0548**	**0.0179**	0.0004	0.0006
pm*(Urban)	0.0050	0.0119	*0.0251*	*0.0140*	**0.0341**	**0.0108**
inc or gsp	**0.0933**	**0.0146**	**0.0545**	**0.0157**	**0.0985**	**0.0146**
adrm	**-0.0162**	**0.0063**	**-0.0198**	**0.0068**	**-0.0128**	**0.0063**
popratio	**0.1476**	**0.0363**	**0.0930**	**0.0385**	**0.0869**	**0.0439**
Urban	**-0.0447**	**0.0207**	-0.0304	0.0224	*-0.0371*	*0.0207*
Railpop	0.0021	0.0082	0.0064	0.0089	-0.0017	0.0082
D7479	**-0.0437**	**0.0035**	**-0.0435**	**0.0036**	**-0.0438**	**0.0036**
Trend	0.0003	0.0003	0.0006	0.0004	-0.0002	0.0003
constant	**2.1014**	**0.1274**	**2.0263**	**0.1466**	**2.0220**	**0.1357**
rho	**-0.0954**	**0.0235**	**-0.1057**	**0.0295**	**-0.0982**	**0.0237**
No. observations	1,785		1,785		1,785	
Adjusted R-squared	0.9805		0.9807		0.9804	
S.E. of regression	0.0318		0.0316		0.0319	
Durbin-Watson stat	1.9196		1.9189		1.9213	
Sum squared resid	1.6861		1.6729		1.7001	

Notes: Bold or italic type indicates the coefficient is statistically significant at the 5% or 10% level, respectively.
In the third equation, the variable *inc*, not *gsp*, is used by itself.
Estimates of fixed effects coefficients (one for each state except Wyoming) are not shown.

Table 6. Rebound Effect with Alternative Specification of Income

	Estimated Using Personal Income		Estimated Using Gross State Product		Estimated Using Time Trend	
	Short Run	Long Run	Short Run	Long Run	Short Run	Long Run
Elasticity of VMT with respect to fuel cost per mile: (a)						
At sample average	-0.0526	-0.2553	-0.0549	-0.2789	-0.0603	-0.3038
Calif. 1997-2001	-0.0220	-0.1133	-0.0292	-0.1544	-0.0456	-0.2328

Tables 5 and 6 also shows a model in which the effect of income on the rebound effect is replaced by letting a time trend, instead of income, affect the relationship between cost per mile (pm) and usage. In this model, inc is still used as a variable by itself, its coefficient being the income elasticity. This version of the model shows a very small and statistically insignificant coefficient for the interaction between time trend and pm. Some other coefficients of the usage equation are affected, including those of adults per road mile (adrm) and population per adult (popratio). The interaction between cost per mile and urbanization shows the biggest change: it is now statistically significant at a 10 percent significance level, and it is large enough to have a moderate effect on the rebound effect. We regard this model as less satisfactory because the time trend is just a mask for unknown effects. However it fits nearly as well as the preferred model and demonstrates that the effect of income on the rebound effect is somewhat tenuously estimated, making it difficult to discriminate between hypotheses based on income and those based on urbanization.

5.4   Caveats

Despite the generally good performance of our equation system, there are many caveats that need to be considered. First, as discussed earlier, there are known problems with the VMT data collected by the US Federal Highway Administration. We have no reason to think that these problems bias the results one way or the other, but better data would add considerably to our confidence in results of this methodology.

Second, our estimates, like those of most previous studies, rely on theory that requires people to react to any change in cost per mile the same way, whether it is caused by variations in fuel prices or in fuel efficiency. As we stressed earlier, there is more variation over time in fuel prices than there is in fuel efficiency, so this theoretical reliance is critical. Our methodology allows us to test for whether fuel intensity (fint) exhibits an independent influence on vehicle usage by simply decomposing the composite variable for price per mile. In fact, in log form the composition is extremely simple: $pm = pf + fint$. The test therefore consists of entering pf and fint separately in the equation instead of combined into pm —both for the variable itself and for its interaction with inc and Urban. We then see whether each of the three pairs of coefficients (involving pf and fint) are equal. In contrast to some earlier studies using different data, such as Greene et al. (1999) and Schimek (1996), we found they are not. The coefficient of pf is very similar to that on pm, but that on fint is small and statistically insignificant. In other words, we cannot prove that there is any rebound effect resulting from stricter fuel efficiency regulations; in the absence of theory, we would have to conclude that fuel price has the expected effect but fuel intensity (or its reciprocal, fuel efficiency) cannot be demonstrated to have a similar effect. However, the model with pf and fint entered separately does not perform very well: the coefficients of interaction terms change greatly and implausibly. Thus we conclude that the best estimate of the rebound effect is attained by using the theoretically justified equating of the effects of fuel price and fuel intensity.

Another way to look at this issue is that few if any observers would expect the influence of fuel intensity to exceed that of fuel price. Therefore, measuring the elasticity of travel with respect to fuel price places an upper bound on the elasticity with respect to fuel intensity, which is the rebound effect. Since we obtain a rather low rebound effect, at least when projected into future years, it seems safest to take this upper bound as our best estimate. We should add that we are skeptical of theories that rely on consumers' lack of awareness of fuel costs. Motorists fill their gas tanks regularly and can easily see the effect on their household expenses, unlike cases such as electrical appliances where people are rarely reminded of what it is costing them.

A third caveat is that we found the role of fuel price in determining fuel efficiency (the model in Table 3) to be sensitive to how the cafe variable is defined. We tried equations with additional variables, including lagged values, in the prediction equation shown in Appendix B for the short time period 1966-78. The time pattern exhibited by the cafe variable was quite different, and the influence of both cafe and fuel price on fint in the structural model (Table 3) diminished to statistical insignificance. However, we determined that the richer specification was unreliable because it was over-fitting the data: coefficients on a variable and its lag were in several instances large and opposite in sign, and the predicted desired fuel intensity showed

implausible oscillations over time. Therefore, we believe the current specification is the most suitable one given the short time period over which we can observe pre-CAFE behavior.

6.   Application of Models: The Rebound Effect in California

6.1   Static Rebound Effects Over Time

In Section 5, we presented estimates of the rebound effect and other elasticities for the US averaged over the sample period 1966-2001, and for California averaged over 1997-2001. The main factor causing them to differ is high income in the latter case.

What can we say about regulations that might go into effect toward the end of this decade? Such projections are hazardous because projected real incomes would become higher than any observed in our data. Therefore any such projection involves extrapolation, which means that even minor deviations of the specific assumptions of the model from actual behavior can cause major discrepancies.

We feel confident in stating that our evidence points to a relevant rebound effect for such regulations that is well below the amount estimated for the US average in our sample, and probably even below that estimated for California in recent years. However we feel that a simple insertion of projected average incomes into our model is too risky, because the influence of the logarithm of real income on the rebound effect is linear in our model and it is unlikely that such a specification can be correct for the large variations in real income that such projections involve. (We did try other functional forms for this influence, but it is too subtle an effect for us to be able to distinguish them from each other statistically.)

In fact, the rebound effect as estimated here becomes zero at incomes that would be reached in projections of 15 years or more. From the coefficients of pm and pm*inc in Table 1, we can see that the rebound effect is zero when the logarithm of real income per capita differs from the mean in our sample by $0.0521/0.0837 = 0.622$ —actually even a little lower due to the small effect of urbanization. Given the mean value of inc in our sample, this occurs at a real income (1987 dollars) of $\exp(0.622+ 9.562) = \$26,490$, or \$41,950 in 2002 dollars. According to projections for California used in Energy Commission modeling, this will occur in approximately 2017. (Note again that this calculation does not involve the coefficient of inc itself, but rather its interaction with fuel cost per mile.)

A more cautious approach is to assume that the rebound effect, rather than declining to zero linearly with log income, will approach zero only gradually. We therefore adopt the following procedure to approximate our logarithmic adjustment with income into an exponential declining one, in such a way as to cause minimal approximation error within the heart of our sample.

(1) First, we define a partial coefficient of pm to account for the income interaction with pm. Using the estimates in Table 1, this is

$$\overline{\beta}_1^m = -0.0518 + 0.0890 * inc . \tag{41}$$

Note that the values of $-\overline{\beta}_1^m$ are a first approximation to the rebound effect, and that according to (41) the rebound effect is a declining concave function of per capita income (INCOME).

(2) Next, we eliminate the lowest 10 percent and the highest 10 percent of those 1,836 observations of $-\overline{\beta}_1^m$; this eliminates extreme values, including those few cases where $-\overline{\beta}_1^m$ is negative. Using the remaining 1,468 data points, we estimate a declining exponential relationship with income to approximate (41) within the range of most of our data:

$$-\overline{\beta}_1^m \approx b_0 \exp(b_1 \cdot INCOME) \tag{42}$$

where INCOME is the unnormalized value of per capita income in \$1000s of 1987 dollars (not in log form). We estimate this equation by taking the logarithm of each side, adding a random error term, and applying ordinary least squares. Equation (42) differs from (41) in requiring that the influence of income on the rebound effect diminish eventually to zero, as one would expect theoretically.

(3) We then apply the equation (42) to projections of future real per capita income for California. Those projections for real per capita income are the values used by the California Energy Commission in its modeling, obtained from Commission staff, approximated by a constant growth rate of 1.64 percent per year starting in year 2000.

(4) For each future projected value of $-\overline{\beta}_1^m$, we make the very small adjustment for the difference between California's urbanization level (0.9763) and its mean value in our sample (0.7129). (California's urbanization level lacks a clear trend over the last three decades so we take it to be constant at its average value for 1997-2001). This gives us the predicted elasticity of usage with respect to cost per mile in our structural model (8), which here we write as:

$$\hat{\beta}_1^m = \overline{\beta}_1^m + 0.0050 \cdot (0.9763 - 0.7129) = \overline{\beta}_1^m + 0.0013 . \tag{43}$$

(5) Finally, with these as the projected future values for β_1^m, we use (37) to project the short- and long-run rebound effects.

The effects of these adjustments are portrayed in Figure 3. The line labeled "logarithmic equation" portrays relationship of equation (41), which is linear in log income and therefore convex in income itself. The dashed lines on either side of it show the 95% confidence bands resulting from uncertainty in the two estimated coefficients in (41). Note that even for incomes so high that (41) predicts negative income effects, the 95% confidence bands include positive values; therefore, these negative values may be viewed simply as the outcome of statistical uncertainty in the model, not as indicating fundamental problems with it. The line marked "exponential fit to 80% of the data" is the fitted rebound effect using (42). It also lies within the 95% confidence bands. Therefore, we view our procedure as a way of choosing among the possible projected values only those that meet the theoretical expectation that the rebound effect should remain positive.

The results of this five-step procedure are shown in the first four columns of Table 7. They portray a long-run rebound effect that starts at 11.2% in 2000 and declines to 7.6% in 2009, 4.3% in 2020, and 2.25% in 2030. The short-run rebound effect (applying in a one-year period) is approximately one-fourth of this value. The results shown for 2000-2001 do not precisely match those shown earlier for California 1997-2001 (in the second row of Table 4) because these are based on a fitted value for the rebound effect, from equation (42), instead of the value calculated for this level of income from directly from the estimated equations (35).

6.2   Dynamic Rebound Effects

Equations (37a,b) describe two out of many dynamic projections that can be made from the structural model in (35). Namely, they describe a scenario in which all coefficients remain fixed and a small permanent change is made in fuel intensity fint, starting in year t. Equation (37a) describes the effect of this on travel in year t, while (37b) describes the ultimate limiting effect after many years. But analysts may be interested in several other dynamic features: (a) they may want to know the effect year by year, not just in year t and in much later years; (b) the coefficients may not remain constant, particularly the coefficient of pm which, in our system, depends on income and urbanization; (c) the scenario of interest may involve a gradual change in fuel intensity rather than a one-time permanent change.

Figure 3. Procedure to Project Rebound Effect
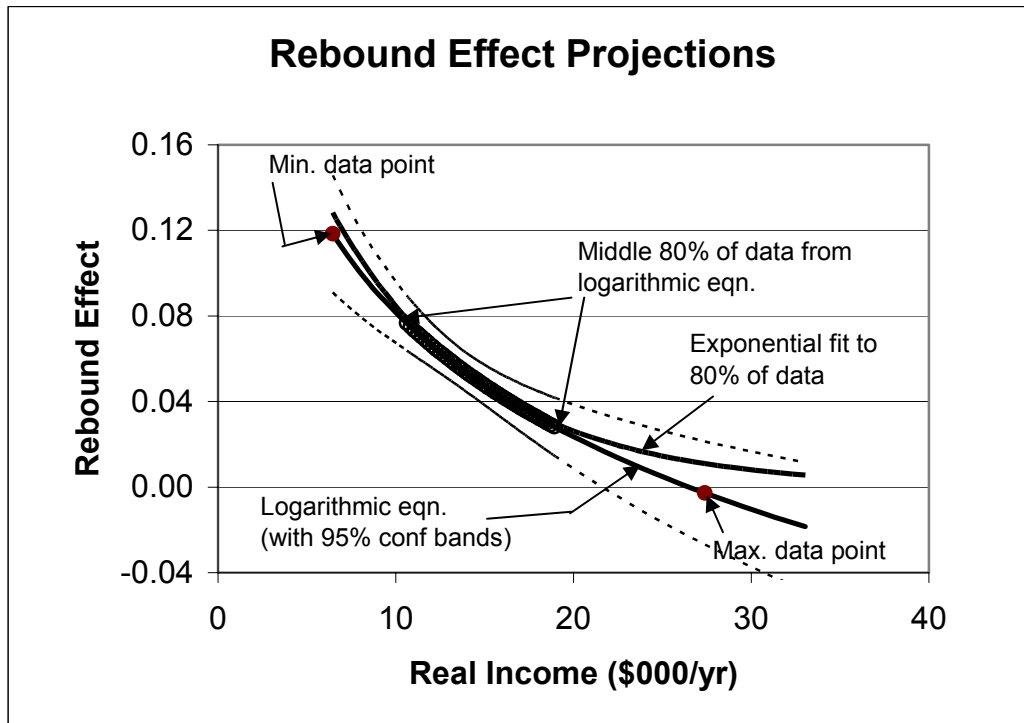
For example, suppose we have a permanent one-time change in fuel intensity, which we assume to be known from some other model, and we want to account for features (a) and (b) above. They work in opposite directions over time: (a) as time goes on the relevant rebound effect gradually changes from its short-run value to its considerably larger long-run value; but (b) rising incomes make both the short-run and long-run rebounds smaller, as shown in Table 7, columns 3 and 4. We call the net effect of these two forces the "dynamic rebound effect."

Using our equation system (35), the dynamic rebound effect, as well as more complex dynamic scenarios, can be simulated by successive application one year at a time, starting in the year the policy goes into effect. In particular, we need the first two of those equations, since we assume here that the outcome of the third equation, namely the change in fuel intensity, Δ*fint*, is known from some other modeling exercise. We will also assume that changes in new-vehicle prices are predicted by an outside model. This is compatible with the use of our model for the kinds of policy exercises typically faced by a regulator considering mandated changes in vehicle technologies in order to reduce greenhouse gas emissions.

These two equations describe the interactive ramifications for vehicle travel and vehicle stock, which affect each other as well as both being affected by the changes in pm and pv that could arise from a greenhouse gas regulation. It is simplest to rewrite these two equations in terms of changes, since then most the coefficients of unchanged variables are eliminated. Recalling that pm≡pf+fint and letting Δ denote the difference in a variable caused by the changes considered by the scenario under consideration (relative to a defined base case), these two equations are:

Table 7. Projected Rebound Effects, California 2000-2030

Year	Income ($1000s)	-------Rebound Effect (%)-------			---------Phase-in Scenario--------- % change in:		
		Short Run	Long Run	Dynamic (start 2009)	new-car E (assumed)	fleet E (assumed)	VMT (calculated)
2000	20.202	2.53	11.51	0	0	0	0
2001	20.536	2.43	11.06	0	0	0	0
2002	20.875	2.33	10.62	0	0	0	0
2003	21.220	2.24	10.19	0	0	0	0
2004	21.571	2.14	9.77	0	0	0	0
2005	21.928	2.05	9.36	0	0	0	0
2006	22.291	1.96	8.96	0	0	0	0
2007	22.659	1.88	8.57	0	0	0	0
2008	23.034	1.79	8.19	0	0	0	0
2009	23.415	1.71	7.82	1.70	4.17	0.42	0.01
2010	23.802	1.63	7.46	2.99	8.33	1.25	0.03
2011	24.195	1.55	7.11	3.95	12.50	2.50	0.06
2012	24.596	1.48	6.77	4.65	16.67	4.17	0.11
2013	25.002	1.41	6.45	5.16	20.83	6.25	0.17
2014	25.416	1.34	6.13	5.50	25.00	8.75	0.25
2015	25.836	1.27	5.82	5.72	25.00	11.25	0.34
2016	26.263	1.20	5.52	5.84	25.00	13.75	0.43
2017	26.697	1.14	5.24	5.89	25.00	16.25	0.52
2018	27.139	1.08	4.96	5.87	25.00	18.75	0.61
2019	27.588	1.02	4.69	5.81	25.00	20.83	0.69
2020	28.044	0.96	4.43	5.72	25.00	22.50	0.76
2021	28.507	0.91	4.18	5.59	25.00	23.75	0.82
2022	28.979	0.85	3.94	5.45	25.00	24.58	0.86
2023	29.458	0.80	3.71	5.29	25.00	25.00	0.89
2024	29.945	0.75	3.49	5.13	25.00	25.00	0.90
2025	30.440	0.70	3.28	4.95	25.00	25.00	0.91
2026	30.944	0.66	3.07	4.78	25.00	25.00	0.90
2027	31.455	0.62	2.88	4.60	25.00	25.00	0.89
2028	31.975	0.57	2.69	4.42	25.00	25.00	0.87
2029	32.504	0.54	2.51	4.24	25.00	25.00	0.85
2030	33.042	0.50	2.34	4.06	25.00	25.00	0.83

Note:

Estimating equation: BetaS = b0 * exp(b1xIncome) where

    b0= 0.2761

    b1= -0.1167

    Income = real 1987 per capita income in $1000s

$$\Delta(vma)_t = \alpha^m \Delta(vma)_{t-1} + \alpha^{mv} \Delta(vehstock)_t + \beta_1^m(t)\Delta(\text{fint})_t$$
$$\Delta(vehstock)_t = \alpha^v \Delta(vehstock)_{t-1} + \alpha^{vm} \Delta(vma)_t + \beta_1^v \Delta(pv)_t + \beta_2^v \Delta(\text{fint})_t$$

(44)

where we have written the coefficient β_1^m as a function of time to remind us that it depends on real income and urbanization, which may change through time. These equations can be solved simultaneously to obtain:

$$\Delta(vma)_t = \frac{\Delta(vma1)_t + \alpha^{mv} \Delta(vehstock1)_t}{1 - \alpha^{mv} \alpha^{vm}}$$

(45a)

$$\Delta(vehstock)_t = \Delta(vehstock1)_1 + \alpha^{vm} \Delta(vma)_t$$

(45b)

where we have defined

$$\Delta(vma1)_t = \alpha_m \Delta(vma)_{t-1} + \beta_1^m(t)\Delta(\text{fint})_t$$
$$\Delta(vehstock1)_t = \alpha^v \Delta(vehstock)_{t-1} + \beta_1^v \Delta(pv)_t + \beta_2^v \Delta(\text{fint})_t$$

(45c)

Equation (45a) is a recursion formula, which can be computed at any time t given the values at time t-1. Therefore we can use it to project *vma* into the future, given a known change at a specific point in time; then (45b) gives us the projected value of *vehstock*. This system can be used to project the results of changes in *fint* or *pv*.

We can see that (44) and (45) are compatible with the short-run and long-run rebound effects already derived in equations (37). Suppose that $\Delta(pv)_t$=0 and that and $\Delta(\text{fint})_t$ is a constant Δ*fint* starting in year 1. To compute the short-run rebound from (45), notice that Δ*fint* affects Δ*vma* through two channels in the numerator of (45a): via Δ*vma*1 (with coefficient $\beta_1^m(1)$) and via Δ*vehstock*1 (with coefficient β_2^v), both divided by (1-$\alpha^{mv}\alpha^{vm}$), thus giving as a net result the right-hand side of (37a). Formally, (37a) is the partial derivative of (45a) with respect to Δ*fint*.

To compute the long-run rebound, we set $\beta_1^m(t)$ equal to a constant β_1^m and look for the solution to (44) in which both *vma* and *vehstock* are constant. Calling these constant values $\Delta(vma)_{LR}$ and $\Delta(vehstock)_{LR,}$ we seek the solution to:

$$\Delta(vma)_{LR} = \alpha_m \Delta(vma)_{LR} + \alpha^{mv} \Delta(vehstock)_{LR} + \beta_1^m(t)\Delta(\text{fint})$$
$$\Delta(vehstock)_{LR} = \alpha^v \Delta(vehstock)_{LR} + \alpha^{vm} \Delta(vma)_{LR} + \beta_2^v \Delta(\text{fint})$$

This can be written as a matrix equation with the solution:

$$y = (I - A)^{-1} \beta \cdot \Delta(\text{fint})$$

(46)

where $y=(\Delta(vma)_{LR}, \Delta(vehstock)_{LR})'$, $\beta=(\beta_1^m, \beta_2^v)'$, I is the 2x2 identity matrix, and

$$A = \begin{pmatrix} \alpha^m & \alpha^{mv} \\ \alpha^{vm} & \alpha^v \end{pmatrix}.$$

Inverting the matrix (I-A) and writing out (46) gives the following result:

$$\begin{pmatrix} \Delta(vma)_{LR} \\ \Delta(vehstock)_{LR} \end{pmatrix} = \frac{1}{D} \begin{pmatrix} 1-\alpha^v & -\alpha^{mv} \\ -\alpha^{vm} & 1-\alpha^m \end{pmatrix} \cdot \begin{pmatrix} \beta_1^m \\ \beta_2^v \end{pmatrix} \cdot \Delta(fint) \qquad (47)$$

where *D*, the determinant of matrix (*I-A*), is given by the denominator of equation (37b). The first component of (47), divided by $\Delta(fint)$, gives the long-run rebound effect and is identical to (37b).

We next use the dynamic system (45) to illustrate the properties of the estimated system for two scenarios. The first, analogous to how the long-run rebound effect is defined, is a permanent 1.0% increase in fuel efficiency. We take this to be in year 2009 for illustration. This permanent change in fuel efficiency will cause an increase in *vma* in the first year (2009) equal to the short-term rebound effect in that year. Subsequently the change in *vma* will rise as we gradually move from the short-run to the long-run effect; but eventually it will fall as rising incomes reduce the rebound effect. The result, expressed as a percentage of the permanent change in fuel intensity, is what we call the "dynamic rebound effect;" it is shown in the fifth column of Table 7. The dynamic rebound effect follows the pattern just described: in the first year it is 1.66%, the same as the short-run rebound effect (within rounding error); it increases to a value of 5.69% in 2016, then gradually declines.

Like the usual rebound effect, the dynamic rebound effect applies to a one-time permanent change in fuel intensity of all vehicles. However, in practice the entire fleet would not change in a single year. We therefore also project a second scenario, in which new-vehicle regulations are introduced gradually, and these new vehicles only gradually enter the vehicle fleet. Specifically, starting in 2009, new cars are subject to greenhouse-gas emissions regulations which cause their fuel efficiency to rise steadily from no improvement to a 25% improvement starting in 2014, at which point the 25% improvement becomes permanent. Second, one-tenth of the fleet in any year consists of new cars that year, with the oldest one-tenth being retired, so that the entire fleet is replaced within 10 years. These assumed fuel efficiency changes are shown in the sixth and seventh columns of Table 7. These assumptions of course only approximate the actual path that California's regulations might take, but they illustrate the nature of the calculations that can be done on any such path.

The results of this second scenario are shown in the last column of the table. They portray extremely small changes in VMT, less than one percent in all years. These changes reach a peak of 0.88 percent in years 2024-2026, then decline. The small size of the projected VMT changes results from the combination of many factors. First, the rebound effect is estimated to be not terribly large to begin with as a percentage of the fractional change in fuel efficiency, about 24% for our sample average in the long run (Table 4, first row, second entry). Second, it is less than one-fourth this size in the short run (Table 4, first row, first entry). Third, it is smaller still in

California in recent years because of its higher incomes (Table 4, second row). Fourth, even using the exponentially declining projections, which are more conservative than our estimated equation itself in terms of how quickly the rebound effect declines, it becomes smaller still in by 2009 due to the anticipated higher per capita income then. Fifth, the rebound effect applies only to the portion of the fleet to which the new regulations were applied; by the time all vehicles have been subjected to the full effects of the regulations (2023 in the scenario), the rebound effect has been further reduced by continuing income growth to less than half its 2009 value. It is through all these factors that a 25 percent increase in fuel efficiency results in increased VMT of less than ½ percent in the early years and only 0.88 percent at its maximum.

6.3 Factors Offsetting the Rebound Effect: Congestion and New Car Prices

Our equation system can be used to simulate two factors that potentially would reduce the rebound effect. First, if the projected increases in VMT cause congestion to increase, this would raise the time costs of driving. This would probably cause some decrease in amount of driving that would not be measured by our system because such time costs are not in the model.

Second, greenhouse gas regulations that cause manufacturers to raise fuel efficiency are likely to increase the cost of manufacturing vehicles and therefore the price of new vehicles. This would cause some reduction in vehicle stock and, according to our findings, that in turn would reduce the amount of driving. In principle, our equation system could be used to estimate such an effect. The elasticities shown in the second panel of Table 4 tell us what to expect from such a calculation. These estimated elasticities are very small, amounting to a long-run decrease in travel of 0.1 percent for every 10 percent rise in new-vehicle prices. However, it should be cautioned that this measurement is based on the statistically insignificant coefficient of pv (new-vehicle price) in the vehicle stock equation. There was not a great deal of variation in the price of new vehicles over the 36 years of our sample, and there was none across states in our data set because we could not find a price index for individual states. Therefore, the most reliable conclusion would be that the price elasticity of new car purchases is not measured well by these data and therefore the ultimate effect of changes in new car prices on amount of driving is uncertain.

7. Conclusions

Using a cross-sectional time series of the 50 US states plus District of Columbia, over a 36-year period, we have estimate equations for motor-vehicle travel demand and for the choice of fuel efficiency. In doing so we produce estimates of the rebound effect, as well as many other elasticities. The equations are based on a more complete theory of what is meant by the rebound effect than has been the case in earlier work. The theory both clarifies how to interpret past empirical results and guides econometric specification.

We argue that accounting for the endogeneity of fuel efficiency is necessary when calculating the per-mile fuel cost elasticity of VMT demand. Doing so substantially reduces the estimated rebound effect. It seems likely that many previous estimates, on the order of 10-20% for short run and higher for long run, would be reduced by around 30 percent if this endogeneity were controlled for. In addition, a better measure of the effects of the federal CAFE standards, in

place since 1978, seems to help stabilize results, which have shown considerable variation in the literature. Our longer time series also enables us to distinguish the effect of a lagged dependent variable (and therefore the difference between short- and long-run effects) from other sources of autocorrelation. Our best estimate of the rebound effect for the US as a whole, over the period 1966-2002, is 5.3% for the short run and 24% for the long run. For California in the recent five-year period 1997-2001, it is 2.2% in the short run and 10.0% in the long run.

Projections to future years are inherently uncertain because California's per capita income is projected to rise to values not seen in any of our sample observations. Nevertheless our best effort at making such a projection yields a long-run rebound effect that declines to about 7.6% in 2009 and 4.3% in 2020.

References

Blair R.D., D.L. Kaserman, and R.C. Tepel, 1984, "The Impact of Improved Mileage on Gasoline Consumption," *Economic Inquiry*, 22(2), April, pp. 209-217.

Bunch, D.S., D. Brownstone, and T.F. Golob, 1996, "A Dynamic Forecasting System For Vehicle Markets With Clean-Fuel Vehicles," in D.A. Hensher, J. King and T.H. Oum, eds., *World Transport Research: Proceedings of 7th World Conference on Transport Research,* (Pergamon, Oxford) Vol. 1, pp. 189-203.

CBO – Congressional Budget Office, 2002, *Reducing Gasoline Consumption: Three Policy Options*, CBO, November 2002.

Davidson, Russell, and James G. MacKinnon, 1993, *Estimation and Inference in Econometrics.* Oxford: Oxford Univ. Press.

Fair, Ray C., 1984, *Specification, Estimation, and Analysis of Macroeconometric Models.* Cambridge, Mass.: Harvard Univ. Press.

Goldberg P. K., 1995, "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry," *Econometrica*, 63(4), July, pp. 891-951.

Goldberg P. K., 1998, The Effects of the Corporate Average Fuel Efficiency Standards, *Journal of Industrial Economics*, 46, 4, 1-33.

Graham D.J. and S. Glaister, 2002, The Demand for Automobile Fuel: A Survey of Elasticities, Journal of Transport Economics and Policy, 36, 1, 1-26.

Greene D.L., 1992, Vehicle Use and Fuel Economy: How Big is the Rebound Effect?, Energy Journal, 13, 1, 117-143.

Greene D.L., J.R. Kahn and R.C. Gibson, 1999, Fuel Economy Rebound Effect for US Households, Energy Journal, 20, 3, 1-31.

Greening L.A. and D.L. Greene, 1997, *Energy Use, Technical Efficiency, and the Rebound Effect: a Review of the Literature*, report to the Office of Policy Analysis and International Affairs, US Department of Energy, Washington D.C.

Greening L.A., D.L. Greene and C. Difiglio, 2000, Energy Efficiency and Consumption – The Rebound Effect – A Survey, *Energy Policy*, 28, 389-401.

Harrington W. and V. McConnell, 2003, *Motor Vehicles and the Environment*, RFF Report, Resources for the Future, Washington D.C.

Harrington W., 2003, Implications of CAFE for Conventional Air Pollutants and Congestion, paper presented to the American Economic Association and the Association of Environmental and Resources Economists, San Diego, Jan.

Haughton J. and S. Sarkar, 1996, Gasoline Tax as a Corrective tax: Estimates for the United States, 1970-1991, *Energy Journal*, 17, 2, 103-126.

Hensher D.A., 1986, Dimensions of Automobile Demand: An Overview of an Australian Research Project, *Environment and Planning A*, 18, pp. 1339-1374.

Hensher D.A., N.C. Smith, F.W. Milthorpe, and P. Barnard, 1992, *Dimensions of Automobile Demand: A Longitutinal Study of Household Automobile Ownership and Use*. North-Holland, Amsterdam.

Hu, P. S. and J. R. Young, 1999, *Summary of Travel Trends: 1995 Nationwide Personal Transportation Survey*. US Federal Highway Administration, Publication No. FHWA-PL-00-006, December.

Jones C.T., 1993, Another Look at U.S. Passenger Vehicle Use and the 'Rebound' Effect from Improved Fuel Efficiency, *Energy Journal*, 14, 4, 99-110.

Kleit A.N., 1990, The Effect of Annual Changes in Automobile Fuel Economy Standards, *Journal of Regulatory Economics*, 2, 151-172.

Kleit A.N., 2002, *Impacts of Long-Range Increases in the Corporate Average Fuel Economy (CAFE) Standard*, AEI-Brookings Joint Center for Regulatory Studies Working Paper 02-10.

Mannering F.L., 1986, A Note on Endogenous Variables in Household Vehicle Utilization Equations, Transportation Research B, 20, 1, 1-6.

NRC, 2002, *Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards*. National Research Council. National Academy Press, Washington, D.C.

NuStats, 2002, *2000-2001 California Statewide Household Travel Survey: Final Report*, California Department of Transportation, June.

Parry I. W. H. and K. A. Small, 2002, Does Britain or the United States Have the Right Gasoline Tax?, Department of Economics Working Paper 01-02-02, University of California, Irvine.

Parry I. W. H. and W. E. Oates, 2000, "Policy Analysis in the Presence of Distorting Taxes." Journal of Policy Analysis and Management 19: 603-614.

Pickrell D. and P. Schimek, 1999, Growth in Motor Vehicle Ownership and Use: Evidence from the Nationwide Personal Transportation Survey, Journal of Transportation and Statistics, 2, 1, 1-17.

Schimek P., 1996, Gasoline and Travel Demand Models Using Time Series and Cross-Section Data from the United States, Transportation Research Record 1558, 83-89.

Thorpe S.G., 1997, Fuel Economy Standards, New Vehicle Sales, and Average Fuel Efficiency, *Journal of Regulatory Economics*, 11, 311-326.

Train K., 1986, *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand*, MIT Press, Cambridge, Mass.

UCLA Anderson Forecast, 2004. http://uclaforecast.com, accessed October 2004.

USDOE, 1979, State Energy Fuel Prices by Major Economic Sector from 1960 through 1977: Preliminary Report and Documentation. US Department of Energy, Energy Information Administration, Washington, D.C.

USDOE, 1996. *Policies and Measures for Reducing Energy Related Greenhouse Gas Emissions: Lessons from Recent Literature*. Report No. DOE/PO-0047, US Department of Energy, Office of Policy and International Affairs, Washington, D.C., July.

USDOE, 2003, *2003 California Gasoline Price Study: Preliminary Findings*, US Department of Energy, Office of Oil and Gas, Washington D.C.

Vincent M. J., M. A. Keyes and M. Reed, 1994, *NPTS Urban Travel Patterns: 1990 Nationwide Personal Transportation Survey (NPTS)*. US Federal Highway Administration, Report No. FHWA-PL-94-018, June.

West, S.E., 2004, Distributional Effects of Alternative Vehicle Pollution Control Policies, *Journal of Public Economics*, 88, 3-4, 735-757.

APPENDIX A: Data Sources and Descriptive Statistics

This appendix lists the variables used in the estimation and their sources, followed by descriptive statistics both for the natural form of the variable and for the form (usually logarithmic) in which we have used it in the equation.

Adult population (18 and over)
> *Definition: midyear population*
> 1966-2001: U.S. Census Bureau (http://www.census.gov)
> > (http://eire.census.gov/popest/archives/1990.php and
> > http://eire.census.gov/popest/data/states/ST-EST2002-ASRO-02.php)
> > (accessed 12/03/2004)

Corporate Average Fuel Economy (CAFE: Mile Per Gallon (MPG))
> 1978-2001: National Highway Traffic Safety Administration (NHTSA), *CAFE Automotive Fuel Economy Program, Annual update 2001*, Table II-6
> (http://www.nhtsa.dot.gov/cars/rules/cafe/FuelEconUpdates/2001/Index.html)
> * Note: The CAFE standards are different from vehicle types. The CAFE standard data
> > used in this study is for passenger cars from "summary of fuel economy performance". (http://www.nhtsa.dot.gov/cars/rules/cafe/CAFEData.htm#)

Consumer price index – all urban consumers (1982-84=100)
> 1966-2001: Bureau of Labor Statistics (BLS), CPI (http://www.bls.gov/cpi/)
> * Note: all monetary variables (gas tax, new passenger vehicle price index, price of
> > gasoline, personal income) are put in real 1987 dollars by first deflating by this CPI
> > and then multiplying by the CPI in year 1987 (divided by 100). The purpose of using
> > 1987 is for ease in replicating Haughton and Sarkar (1996).

Cumulative Population Growth Rate (%)
> Definition: Percentage of population in year t that has been added since 1950.
> 1950: *Statistical Abstract of the United Sates* (SAUS)
> 1960-2001: Bureau of Economic Analysis (BEA) (http://www.bea.doc.gov)
> * Note: Original source for both is from midyear population estimates by U.S. Census
> > Bureau

Family size
> *Definition: average number of people per household*
> 1966-2001: Bureau of Labor Statistics, Current Population Survey (CPS)
> > (http://www.bls.gov/cps/)
> * Note: Data was extracted from a CPS Data CD

Federal Gas Tax (cents per gallon)
> 1966-2001: Federal Highway Administration (FHWA), *Highway Statistics*, Annual
> > Report, Table FE-101A 1A

Highway use of gasoline (including public use) (thousands of gallons)
> 1966-1995: FHWA, *Highway Statistics Summary to 1995*, Table MF-226

1996-2001: FHWA, *Highway Statistics Annual Report*, Table MF-21
* Note: The FHWA estimates highway use of gasoline by subtracting estimated non-highway use from the total use reported by States.

Income per capita ($/year, 1987 dollars)
   Definition: Personal income deflated to 1987, divided by midyear population
   1966-2001: Bureau of Economic Analysis (BEA) (http://www.bea.doc.gov/)
   * Note: Per capita personal income is total personal income divided by total midyear population.
   Alternative measure: Disposable income (similarly deflated and divided by population); available from same web site as above, but only starting 1969; for 1966-68 we interpolated by assuming it bore the same ratio to per capita personal income as existed in the same state for 1969-78.
   Second alternative measure: Gross state product (similarly deflated and divided by population), available starting 1977 from the same source. For 1966-1976 we interpolated by assuming it bore the same ratio to per capita personal income as existed in the same state for 1977-86.

Interest rate (%)
   Definition: national average interest rate for auto loans
   1966-1971: Interpolated using Moody's AAA corporate bond interest rate
   1972-2001: Federal Reserve Systems, Economic Research and Data, Federal Reserve Statistical Release G.19 Consumer Credit
   (http://www.federalreserve.gov/releases/g19/hist/cc_hist_tc.html)
      *Note: We average two different interest rates: that for new-car loans at auto finance companies, and that for commercial banks for 48-month lanes for new car. These two rates are averaged over a year from monthly and quarterly data, respectively.

New Car Price Index
   Definition: price index for U.S. passenger vehicles, city average, not seasonally adjusted, (1982-84=100)
   1966-2001: Bureau of Labor Statistics (BLS), CPI
   * Note: Original index has 1982-84=100; converted to 1987=100 using the Consumer Price Index.

Number of vehicles
   Definition: Number of automobiles and light trucks registered
   1966-1995: FHWA, *Highway Statistics Summary to 1995*, Table MV-201
   1996-2001: FHWA, *Highway Statistics Annual Report*, Table MV-1
   *Note: Trucks include pickups, panels and delivery vans; beginning 1985, personal passenger vans, passenger minivans, and utility-type vehicles are no longer included in automobiles but are included in trucks.

Price of gasoline (cents per gallon)
   1960-1977: U.S. Department of Energy (DOE), *State Energy Fuel prices by Major Economic Sector from 1960 through 1977*, Table B-1, pp. 93-94

1970-2000: Energy Information Administration, *State Energy Data 2000: Price and Expenditure Data*, Table 5

2001: Energy Information Administration, *Petroleum Marketing Annual*, Table A1.

* Note: Data for 1966-1970 were spliced by averaging the two overlapping sets of data.

Public road mileage (miles)

1966-1979: FHWA, *Highway Statistics*, annual editions, Table M-1 (Total rural and municipal mileage)

1980-1995: FHWA, *Highway Statistics Summary to 1995*, Table HM-220

1996-2001: FHWA, *Highway Statistics*, annual editions, Table HM-20

Rail Transit Availability Index (Railpop)

Definition: The fraction of the state's population living in metropolitan statistical areas with a subway or heavy rail transit system

Heavy rail transit system Initial Segment opening year: American Public Transportation Association (APTA) (http://www.apta.com)

(Rail transit dummy for years: 1= rail transit available, 0=otherwise)

Population data for the Metropolitan Statistical Areas (MSA) (1966-2001): SAUS, "Metropolitan Statistics" section (table number changes)

* Note Data for missing years (1969, 1971, 1974, 1979, 1981, 1982, 1989) were interpolated using its state population in which the MSA is included.

Number of Licensed Drivers

1966-1995: FHWA, *Highway Statistics Summary to 1995*, Table DL-201

1996-2001: FHWA, *Highway Statistics*, annual editions, Table DL-1C

Urban Road Mileage (miles)

1966-1979: FHWA, *Highway Statistics*, annual editions, Table M-1 (Total municipal mileage)

1980-1995: FHWA, *Highway Statistics Summary to 1995*, Table HM-220

1996-2001: FHWA, *Highway Statistics*, annual editions, Table HM-20

Urbanization

Definition: Share of total state population living in Metropolitan Statistical Areas (MSAs), with MSAs based on December 2003 definitions

1966-1968: Extrapolated exponentially (i.e. assuming constant annual percentage growth rate) from 1969-79 values

1969-2001: Bureau of Economic Analysis, Regional Economic Accounts (http://www.bea.doc.gov/bea/regional/reis/)

VMT (Vehicle Miles Traveled, , million miles)

1966-1979: FHWA, *Highway Statistics*, annual editions, Table VM-2

1980-1995: FHWA, *Highway Statistics Summary to 1985*, Table VM-202

1996-2001: FHWA, *Highway Statistics*, annual editions, Table VM-2

Table A1. Summary Stastistics for Selected Variables

	Mean	Std. Dev.	Minimum	Maximum
Vma	10,977	2,584	4,737	23,333
Vehstock	1.0026	0.1901	0.4527	1.7433
Fint	0.0615	0.0124	0.0344	0.0919
Pf	108.92	23.52	60.32	194.91
pf	0	0.2032	-0.5696	0.6033
PM	6.8140	2.2754	2.7820	14.2046
pm	0	0.3490	-0.8369	0.7935
ADRM	57.49	68.14	2.50	490.20
INCOME	14,588	3,311	6,448	27,342
inc	0	0.2275	-0.7909	0.6538
DISPINC	12,800	2,764	5,834	24,125
GSP	18,139	6,889	7,763	72,056
gsp	0	0.2929	-0.7996	1.4285
Popratio	1.4225	0.0900	1.2314	1.8616
Urban	0.7129	0.1949	0.2895	1.0000
Railpop	0.0884	0.2073	0.0000	1.0000
Pv	1.0663	0.1972	0.7775	1.4930
Interest	10.83	2.41	7.07	16.49
Licad	0.9085	0.0848	0.6239	1.2286

Note: Units are as described in Section 5.2.
Variable whose names have no capital letters are in natural logarithms; variables whose names are all capital letters are unnormalized.

APPENDIX B: Variable Measuring Strength of CAFE Regulation

Steps in creating the variable

1. We first estimate the reduced-form equation explaining fuel intensity—i.e., the empirical counterpart of the third of equation set (10)—on data only from 1966-1977, with no regulatory variable included (since there was no regulation then). This equation should in principle include all exogenous variables from all three models (including PV for the V equation); we implied it by dropping the variable Railpop, which seemed to have little effect in this short time series. Like our other equations, it also includes one lag of the dependent variable, and allows for fixed effects and autocorrelated errors. It does not include other endogenous variables, either current or lagged; the reason is that, unlike in an instrumental variables regression, our objective is to estimate a predictive model for what fuel intensity would have been in the absence of CAFE regulation and therefore we cannot use information about what actually happened to the endogenous variables. In theory, this equation could include any number of lagged values of independent variables, because they would be present in a complete solution of system (35) for the time path of fint; however on this very short time series it is impractical to estimate so many parameters, especially of variables that are highly correlated as current and lagged values are likely to be. For the same reason of parsimony, we included only a single time trend in this predictive equation.[26] Let us denote this estimating equation by the following reduced-form and simplified variant of the third of equations (35):

$$(fint)_{i,t} = \alpha^{fR} (fint)_{i,t-1} + \beta^{fR} X_{it}^{fR} + u_{it} \tag{B1}$$

where i designates a state, superscript R indicates the reduced form, and XfR denotes the set of all exogenous variables used, including prices, as described above. The results of this estimation are shown in the Appendix. The statistically significant coefficients are that of $(fint)_{t-1}$ with value 0.638, D7479 with value -0.021, and pv with value -0.221. The price of fuel is not statistically significant (t-statistic -1.02) but has the reasonable value of -0.021.

2. The coefficient αfR of the lagged dependent variable is interpreted as arising from the following partial adjustment model:

$$(fint)_{i,t} = (fint)_{i,t-1} + \gamma \cdot \left\{ (fint)_{i,t}^* - (fint)_{i,t-1} \right\} + u_{it} \tag{B2}$$

where $(fint)_{i,t}^*$ denotes a long-run desired value for the logarithm of fuel intensity. That is, users basing decisions in year t desire to shift the vehicle stock toward one with fuel efficiency $(fint)_{i,t}^*$ but they can do so only part way by changing a portion γ of the stock in that year. Thus it is natural to interpret $(fint)_{i,t}^*$ as the target fuel efficiency for new car purchases and γ as the

---

[26] The variable *Trend3* could not be included in any event because it is zero for the time period 1966-77. The variable *Trend2* differs form *Trend* only during the last four years of this period, one of which (1974) already has a separate dummy variable, so it is not surprising that we found it impossible to get precise measures its separate effect.

fraction of the fleet that turns over each year. It is easy to see that (B2) is the same as (35c) with café=0 if we choose $\gamma=1-\alpha f$ and

$$\left(\mathrm{fint}\right)^*_{i,t} = \frac{\alpha^{fm}(vma)_t + \beta^f_1(pf)_t + \beta^f_3 X^f_t}{1-\alpha^f}. \tag{B3}$$

This formula for desired fuel intensity is computed for each state and each year t, not just the years from which the coefficients were estimated.[27]

3. We then form from this the US average desired fuel intensity, averaged the same way as vehicles are averaged under CAFE regulations: namely,

$$(FintUS)^*_t \frac{\sum_i M_{it} \exp\left((fint)^*_{it}\right)}{\sum_i M_{it}} \tag{B4}$$

where Mit is aggregate VMT for state i.

4. Finally, we assume CAFE is binding whenever the desired efficiency $E^*_t \equiv \left(1/FintUS^*_t\right)$ is less than the minimum mandated efficiency, $\overline{E}_t$. The latter is computed as a weighted average of the CAFE standards for light trucks and cars, the weights being current nationwide light truck and car VMT, reduced by 16 percent which is an estimate of the difference between fuel efficiency achieved in real driving and that achieved on the tests used to enforce the CAFE standard.[28] A measure of the strength of CAFE regulation is then

$$R_E \equiv \max\left\{\frac{\overline{E}_t}{E^*_t}, 1\right\}$$

or its logarithm,

$$cafe \equiv \max\left\{\left(\overline{e}_t - e^*_t\right), 0\right\} \tag{B5}$$

where $\overline{e}_t = \ln(\overline{E}_t)$ and $e^*_t = \ln(E^*_t)$.

---

[27] Note that the variable *D7479* is just a dummy variable for year 1974 in this shorter data set. By considering the variable to be a dummy for both 1974 and 1979, we are assuming that the queues at gasoline stations, which appeared in both years due to disruptions in the world market for crude oil, have the same effects in the two years. In fact we found that we could estimate separate coefficients for a 1974 dummy and a 1979 dummy on the full system, and they are nearly equal, giving us confidence in the validity of this assumption.

[28] The factor 16 percent is taken from Harrington (2003).

Estimated Equation for Projecting Desired Fuel Intensity

As described above and in Section 5.2.3, we estimated a reduced form equation to explain the logarithm of fuel intensity, *fint,* as a function of all exogenous variables in the system, using the same approach to autocorrelation and fixed effects as in our structural equations. This equation is estimated for the pre-CAFE period 1966-1977. It does not have high precision due to the short time period, and some of the variables may be operating through more than one channel. The results are shown in Table B1 below. We use the pre-CAFE period because we want to estimate what people chose to do in the absence of CAFE regulation. We use a reduced form because we want to use it to project into the regulated period and so cannot use any variables that are affected by regulations themselves. We include the lagged dependent variable because that will be used to infer the parameter of the partial adjustment model, equation (B2); values of the lagged dependent variable are not used to form the projected desired fuel efficiency.

Because the equation is reduced form, there is no endogeneity to correct. Therefore the equation is estimated by least squares with correction for autocorrelation.

Table B1. Fuel Intensity Equation:
**Reduced Form Estimated on 1966-1977 Data**

Variable	Coefficient	Std. Error
fint(t-1)	0.6386	0.0443
pf	-0.0209	0.0204
Inc	0.0169	0.0288
adrm	0.0363	0.0273
popratio	0.0852	0.0910
Urban	-0.1974	0.2328
D7479	-0.0213	0.0060
Trend	-0.0097	0.0024
pv	-0.2209	0.0798
Interest	0.0213	0.0298
licad	0.02605	0.0262
constant	-0.9822	0.3584
Rho	-0.1241	0.0625
No. of observations		510
Adjusted R-squared		0.8967
S.E. of regression		0.0253
Sum squared resid		0.2858
Durbin-Watson stat		1.9975

Note: 50 constants for individual states are not shown.

Results.

Figure 2 shows the results of this procedure. It compares our estimate of desired nationwide fuel efficiency (E^*) with the de facto standard ($\overline{E}$). We see that the desired efficiency of new passenger vehicles was mildly increasing over much of our time period, especially 1975-1978 and 1984-1997, with one-year upticks in 1974 and 1979 due to queues at gasoline stations and small downturns in 1986 and 1998-99 due to decreases in fuel prices.[29] The CAFE standard exhibited a very different pattern, rising rapidly from 1978-1984 and then flattening out. The variable cafe is zero until 1978, after which it is the logarithm of the ratio of these two values. We can see that by this definition, the CAFE standard has been binding throughout its time of application, but that its tightness rose dramatically during its first six years and then gradually diminished until it is just barely binding in 2001. This pattern is obviously quite different from either a trend starting at 1978, or the CAFE standard itself, both of which have been used as a variable in VMT equations by other researchers.

Figure B1. Desired and Mandated Fuel Efficiencies

---
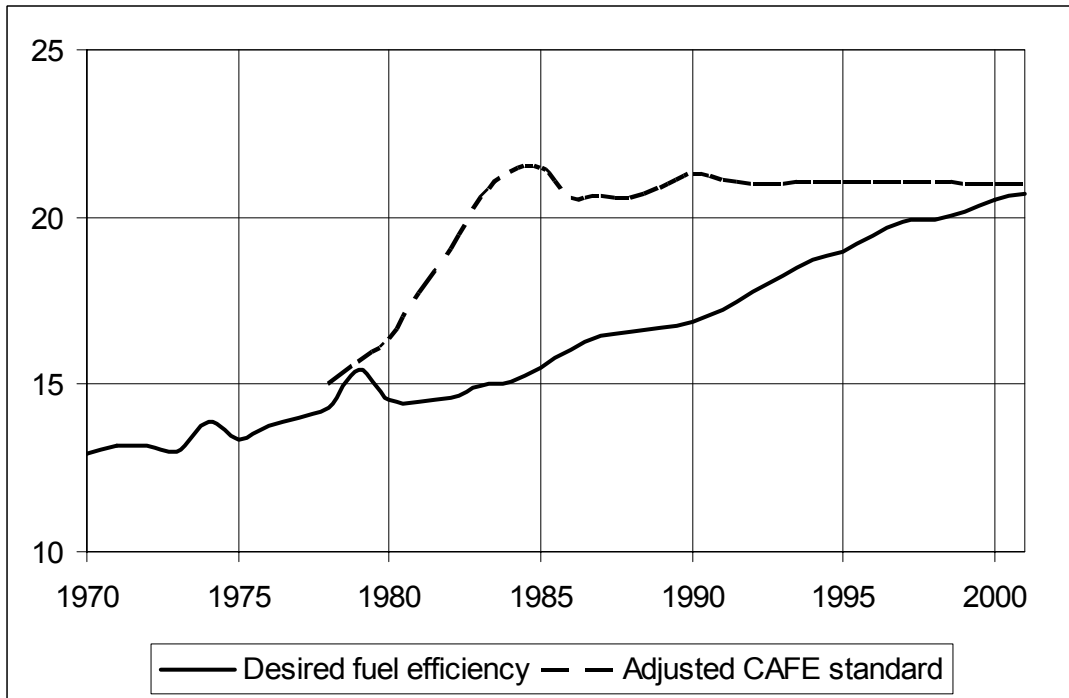
[29] The uptick in 1979 results from our assumption that the gasoline queues in 1979 would have the same effect on desired efficiency as those in 1974, which are captured by the 1974 dummy variable in the equation for fuel intensity fit on 1966-1977 data.

APPENDIX C: Compendium of Reviewers' Comments

The following is a summary of comments (exclusive of typographical and spelling errors) received from reviewers of the report. Each is followed in square brackets by a description of actions taken by the authors in response, indexed to the current page numbering.

California Energy Commission

1. *We would expect income to have a positive impact on VMT, since travel per household tends to rise with disposable income. In addition, since Small uses a macro sample, income per capita is a general economic indicator, so that rising income per capita means better economic conditions, more jobs, and therefore more commute travel. However, the coefficient estimated by Small for income in the usage equation (Table 1) is negative, -0.074. At the same time, the coefficient for the income/fuel cost per mile interaction is positive, 0.089. The total impact of income on driving is given by the net result of these two coefficients. Given that we expect a positive impact from income, a negative coefficient for income by itself means that some of this expected positive impact is being captured by the income/fuel cost per mile interaction.*

Response: This comment was based on an inadequate explanation in the draft final report of how the interacted variables in Table 1 must be used to calculate the net effect of income. The total impact of income on driving is actually in the expected direction. To make this easier to see in the report, the equations were re-estimated with the two interacted variables, income and fuel cost per mile, both expressed as deviations from their mean value. This does not affect any of the estimated effects but makes them more transported because now the reported coefficient of income is equal to the income elasticity measured at the mean values for the sample, as explained in new text on p. 38 and depicted in the new Figure 2.

2. *If one attempts to predict a VMT change from one year to the next using the formulation given in the July 9 draft, a misleading conclusion could result. As a simple example, suppose that efficiency increases (pm decreases) from year t to year t+1 while per capita income also increases, and nothing else changes. The equation for VMT given in the July 9 draft would yield a total increase in VMT (from the efficiency and income impacts) smaller than the apparent impact of the decrease in* pm.

Response: We now illustrate in detail in section 6.2 how to use the results in dynamic simulations, in order to forestall this kind of improper use of the results.

3. *It is worrisome that predicted fuel intensity did not yield a significant coefficient when entered in the VMT equation by itself (Section 5.4).*

Response: We discuss this concern in the second and third paragraphs of Section 5.4, and also note it in the Executive Summary in the second full paragraph of p. ix. Since no one argues that people would pay *more* attention to fuel efficiency than to fuel price, our theoretical restriction equating the effects of these two components of fuel cost per mile puts an upper bound on the rebound effect.

4. *Can you also provide summary statistics?*

Response: These are now provided in Appendix A, Table A1.

NERA

1. *As documented in reports prepared by NERA and Robert Crawford, the results of the UCI study are inaccurate because of mistakes made in formulating the models used in the study – when those mistakes are corrected, the magnitude of the rebound effect calculated using the UCI methodology is essentially the same as that found elsewhere in the literature. The ISOR also states that the travel demand models used by the Southern California Association of Governments and the Bay Area Metropolitan Transportation Commission show no significant rebound effect. The following section of this report contains an explanation of why travel demand models are not capable of estimating the rebound effect. (Sierra Research, Review of the August 2004 Proposed CARB Regulations to Control Greenhouse Gas Emissions from Motor Vehicles: Cost Effectiveness for the Vehicle Owner or Operator, page 21)*

Response:

The reports referred to make three main assertions about our study of the rebound effect:

> (1) The interaction variable, cost per mile times income, is said to be very highly correlated with cost per mile, making it impossible to measure their effects separately. This would be important because it is the coefficient of the interaction variable that measures how the rebound effect changes with real per capita income, and this has a major effect on the projected results in California in 2009 and beyond.

> (2) The problem noted in (1) is said to manifest itself in an unrealistic estimate of the income elasticity, i.e. the parameter measuring how much VMT increase with income. Mr. Crawford asserts that our income elasticity is negative, whereas theory and other studies would lead one to expect it is positive.

> (3) California's high per capita income is said to be increasingly offset by higher living cost. Mr. Crawford prefers a model where income is measured as disposable income (which is after-tax) divided by a cost of living adjustment that he computes using data from ACCRA.

> (4) Gross State Product (GSP) is a better measure of income that per capita personal income.

The first point is based on a misunderstanding of how correlations affect statistical results. The second is based on a mistake in computing the income effect from our results. The third is a dubious use of cost-of-living figures that apply only to metropolitan areas. The fourth is a

legitimate alternative, but on balance we think it is inferior. We explain these in more detail as follows.

(1) Correlation between variables
As long as a statistical model includes a constant term, as ours does, the effect of variables on statistical results depends on their variation within the sample, not their absolute values. For example, a common practice is to "normalize" a variable by subtracting from its values the average value in the entire sample. When two variables are multiplied to create an interaction variable, as in our model, each of them can be similarly normalized by subtracting its mean. In our draft report, we did this for income but not for the variables it is interacted with. This affects the measured correlations between variables, but not the coefficients, model predictions, or standard deviations of coefficients. The choice as to whether to normalize a variable or not is made solely on the basis of convenience in presentation.

In the revised report we have simplified the presentation by normalizing cost per mile (*pm*) as well as income, greatly reducing the measured correlation between variables so that it is more apparent that such correlation does not interfere with statistical properties of the model.

(2) Income Elasticity
The assertion about the income elasticity of VMT in our model is due to a mistaken calculation based on the way the model was presented in the draft report. Specifically, that calculation used only the coefficient of income and neglected to account for the effect of income through the interaction variable between income and *pm*. Doing the calculation correctly would have given the same income elasticity (at mean values of the variables) as is now presented in Table 4. By normalizing variable *pm*, as just described under (1), this is now easier for the reader to understand.

(3) Cost of Living Adjustment
Our monetary numbers, including income and fuel costs, are all stated in real terms, i.e. deflated by a cost of living index. This index is the US nationwide consumer price index. Ideally we would have liked to include variations across states in cost of living. However, no state-wide cost of living indices are available. Furthermore, the differential impact of income across states on driving may not be affected much by the state's cost of living, because it depends on the value of time. These points are now explained on p. 34.

The cost-of-living indices suggested by NERA are unsatisfactory. For one thing, they are computed only every five years. More important, they apply on to metropolitan areas; the variable NERA proposes as an alternative is not a state-wide index, but rather an index averaged over that state's metropolitan areas. Use of this index creates an error in measuring real income that varies in unknown ways according to the state's level of urbanization. Real income would be misstated in mostly rural states relative to mostly urban states, potentially confounding the effects of other variables. For example, the

income in mostly rural states would be understated because income would be deflated by a higher index than the one actually applying to the state as a whole. Since rural states tend to have lower incomes, this bias would exaggerate the differences in real income between rural and urban states, and thereby underestimate the effects of those differences on driving patterns.

Furthermore, if a different cost of living index were used for income, it should be used also to compute real fuel price. That would change the variable "cost per mile," which is critical for measuring the influence of income on the rebound effect.

<u>(4) Measure of Income</u>
We now discuss on p. 43 the effect of using either disposable income or Gross State Product as a replacement for per capita income, and explain why the latter is a less reliable income measure for our purposes. Results for Gross State Product are shown in Table 5.

2. *The procedure to deal with uncomfortable results of negative rebound rates is strange and calls into question the validity of the study. For each year and state in their sample, the authors use the coefficients for price and the price-income interaction term (but not the price-urbanization term) to generate estimates of both short-run and long-run rebound effects. They then throw out "the lowest 5 percent and the highest 5 percent" of their predicted values, thereby eliminating the negative rebound estimates (although this feature is not noted). Next, they regressed the natural logarithm of the predicted rebound effect (truncated to eliminate the negative effects) on income alone. Finally, they used the predicted values from this equation to forecast the rebound effect in California in future years. The net effect of this procedure (not acknowledged by the authors of the study) is to eliminate the uncomfortable results but also to introduce inherent biases in the projected rebound effects. (NERA Economic Consulting, Reviews of Studies Evaluating the Impacts of Motor Vehicle Greenhouse Gas Emissions Regulations in California, page 27).*

Response: The problem addressed is that stated in comment 3 below, and our response to that comment applies here. We adopted the procedure described in this comment in order to be more conservative in projecting future declines in the rebound effect than if we simply projected the point estimates, as well as to avoid projecting negative values. The language on p. 46 has been changed to be sure this reasoning is transparent.

3. *The rebound effect results for some states are not plausible. The results of the Irvine study suggest that some states have negative rebound effects—ie., that drivers there are actually likely to drive less if the cost of driving falls. These results are clearly not sensible, suggesting problems with the underlying model. Moreover, this problem apparently led the authors to develop an ad hoc (and flawed) method to modify their projections of rebound effects in California rather than use estimates from their model. (NERA Economic Consulting, Reviews of Studies Evaluating the Impacts of Motor Vehicle Greenhouse Gas Emissions Regulations in California, page 29).*

Response: This comment refers to the fact that using our model to predict the rebound effect for every state in every year produces some (small) negative values, contrary to theory. This observation does not contradict the validity of the findings. On the contrary, like any statistical study, ours implies a margin of error around any specific predictions. Figure 3 has been added to the final report, showing a 95 percent confidence interval around the projected rebound effect as a function of income. This shows that plausible positive rebound effects are within this confidence interval in each case, as is now pointed out on p. 47.

4. *Key data on income and gasoline price do not reflect state differences in the cost of living. Several of the key data and variables in the Irvine study are unreliable, with ramifications for model estimation. In particular, the income and price variables are not appropriately deflated to take into account differences across states in the cost of living. (NERA Economic Consulting, Reviews of Studies Evaluating the Impacts of Motor Vehicle Greenhouse Gas Emissions Regulations in California, page 29).*

Response: These points are covered in our response to #1 above.

5. *The specification of the trend variables is arbitrary, and alternative specifications are superior. There is no clear conceptual rationale for the authors' choice of trend variables. Alternative specifications are superior. (NERA Economic Consulting, Reviews of Studies Evaluating the Impacts of Motor Vehicle Greenhouse Gas Emissions Regulations in California, page 29).*

Response: This comment refers to the "time trend" variable, which we included in a very standard manner for studies with data covering an extensive time period. We have reworded the report to note that there is a clear theoretical rationale for having three trend variables in the fuel economy equation, due to responses to oil crises in 1973 and 1979; but there is no such rationale for the usage and vehicle stock equations, which is why we do not use them there. In our early experiments, we did not find any important differences that result from allowing three separate trends. Although the NERA commentary indicates a statistical rejection of equality of the hypothesized three trend variables they propose, it does not state whether the resulting coefficients portray a sensible pattern. We found generally that the time trend variables play very little role in our analysis, indicating that most of the important trends are captured in the variables that we include.

6. *Errors in variables and equation specification call the estimation procedure into question. Errors in the data underlying the* licad *variable and in the construction of the* cafe *variable suggest that three-stage least squares is not the best estimator for obtaining estimates of the rebound effect. (NERA Economic Consulting, Reviews of Studies Evaluating the Impacts of Motor Vehicle Greenhouse Gas Emissions Regulations in California, page 29).*

Response: This response refers to properties of the three-stage least squares (3SLS) estimation technique, which tends to cause any data errors in one equation to cause inaccuracies in other

equations. However, precisely because of this, we show results for two alternatives to 3SLS, noted in the report on p. 37. It is easy to see in our tables that the results of interest from two-stage least squares and three-stage least squares are very similar, so it really doesn't matter which one is used. We also note on p. 37 the advantage of 3SLS in achieving greater precision.

7. The VMT equation should be estimated using two-stage least squares rather than three-stage least squares. (NERA Economic Consulting, Reviews of Studies Evaluating the Impacts of Motor Vehicle Greenhouse Gas Emissions Regulations in California, page 40).

Response: Same as #6 above.

8. The income and gasoline price data should be deflated using appropriate state-specific price indices that account for state differences in the cost of living and changes over time. NERA Economic Consulting, Reviews of Studies Evaluating the Impacts of Motor Vehicle Greenhouse Gas Emissions Regulations in California, page 40).

Response: same as #1 above, point (3).

9. Given the arbitrariness of the trend variables, the specification should be modified to reflect the most appropriate set of trend variables.   (NERA Economic Consulting, Reviews of Studies Evaluating the Impacts of Motor Vehicle Greenhouse Gas Emissions Regulations in California, page 40).

Response: same as #5 above.

10. We used data from the American Chamber of Commerce Research Association to develop state-specific cost of living indices. Relying on city and regional CPI data from the BLS, we developed state-specific CPIs for the period from 1996-2001. The Irvine study uses Trend rather than the Trend1, Trend2, and Trend3 to explain both the VMT equation and the vehicle stock equation. Also the two dummy variables D74 and D79 are theoretically superior replacements for the single dummy variable D7479 in the VMT and fuel efficiency equations. We have adjusted the Irvine study estimation to incorporate both of these improvements. These modifications cause the long-run rebound effect for California to nearly triple from 9.3 percent in the Irvine Study to 24 percent or 25 percent under the two revised models. (NERA Economic Consulting, Reviews of Studies Evaluating the Impacts of Motor Vehicle Greenhouse Gas Emissions Regulations in California, pages 41-42).

Response: same as #1 above, point (3).

11. The authors adopt an ad hoc approach to projecting income in order to prevent the projected rebound effects from becoming negative in California in the future, as the model implies already

happens for several states. This ad hoc approach is not defensible and thus we do not present corrected results using this procedure. (NERA Economic Consulting, Reviews of Studies Evaluating the Impacts of Motor Vehicle Greenhouse Gas Emissions Regulations in California, page 43).

Response: same as #2 above.

12. *We have reviewed the data and modeling approach of the Irvine study. Our review identified the following four primary concerns with the Irvine study model: The rebound effect results for some states are not plausible, casting doubts on the model. Key data on income and gasoline price do not reflect state differences in the cost of living. The specification of the trend variables is arbitrary and alternative specifications are superior. Errors in equation specification and key variables call the specific estimation procedure used into question.*

The first of these concerns suggests an underlying problem with the modeling approach, because it generates results that are not plausible. The other three concerns concern specific issues with data and estimation that can be corrected.

As a result of these concerns, we developed three major modifications to the modeling: We re-estimate the model using two-stage least squares rather than three-stage least squares. We deflate the income and gasoline price data using appropriate state-specific price indices that account for state differences in the cost of living and changes over time. We modify the specification to reflect the most appropriate set of trend variables.

These modifications lead to revised estimates of the California rebound effect of 5.3 percent in the short run and 24 percent in the long run, substantially greater than the rebound effect estimates developed in the Irving study. (NERA Economic Consulting, Reviews of Studies Evaluating the Impacts of Motor Vehicle Greenhouse Gas Emissions Regulations in California, pages 43-44).

Response: NERA presents no evidence concerning which of its proposed revisions affects the results in an important way. The many experiments we carried out with alternative specifications makes us believe that the only one that has any substantial effect on the results is the use of metropolitan-area cost-of-living indices to apply to entire states. See our response to #1 above, point (3), for why we believe these cost-of-living indices are inappropriate.

Michael Hanemann

I have focused so far on the consumer's choice of whether and when to buy an automobile, and which model to select. Another important issue is the utilization decision of how much to drive. This has potential environmental as well as economic implications – it affects the consumption of gasoline and, to the extent that there are environmental impacts from the production and consumption of gasoline and/or the driving of automobiles, it affects these as well. The latter have received considerable attention in the economic literature on the effects of the CAFÉ

standards. It has been argued that promoting greater use of cars with a higher mileage per gallon reduces the cost of driving and encourages more driving and more fuel consumption. This is known in the energy literature as the rebound effect. The crucial question is the quantitative magnitude of this effect.

There is a substantial literature on the rebound effect for motor vehicle utilization in the context of the federal CAFÉ standards. This literature is reviewed in some detail by Small and Van Dender (2004). Following their literature review, Small and Van Dender estimate an econometric model consisting of three equations for (the logs of) the vehicle miles traveled in a state per adult of the population, the number of vehicles in a state per adult of the population, and the fuel intensity of the state (measured as the highway gasoline usage in a state divided by the vehicle miles traveled by the adult population). The system of equations allows for simultaneity in the determination of all three variables and it is estimated by three stage least squares, using annual data on the 50 states of the US plus the District of Columbia for the period 1966-2001.The rebound effect – the elasticity of vehicle miles traveled per capita with respect to fuel cost per mile – is estimated to be smaller for California than for the other states, the estimate being 2.0% in the short run and 9.3% in the long run. The estimates imply that, if the operating cost of a motor vehicle in California decreases by 25% in year 2009, the number of miles traveled would increase by 0.17% in 2009 (the short-run effect), by 0.28% in 2010, and by 0.32% in 2020 (staff report Table 11.3-1).

The fuel cost variable used by Small and Van Dender is the price of gasoline divided by mileage per gallon. Like most other authors in the literature on the rebound effect, Small and Van Dender assume that drivers respond to a change in the fuel cost per mile in the same way regardless of whether this is caused by a change in the price of gasoline or change in fuel efficiency (mileage per gallon). However, when Small and Van Dender test this assumption, they find that it does not hold: miles driven responds differently to a change in the price of gasoline than to a change in fuel efficiency. In fact, they cannot reject the hypothesis that changes in fuel efficiency have no effect on miles driven, which would imply a rebound effect of zero.

Based on the information presently available to me, I agree with Small and Van Dender's conclusion that the rebound effect in California is likely to be small if not zero.

While I believe that the rebound effect associated with the regulations is likely to be small if not zero, I also believe that there are serious shortcomings in the data used in the existing literature which make it unlikely that one can obtain a reliable measure of the rebound effect from these data. This criticism is not limited to the data used by Small and Van Dender: it applies to the data used in all the papers that I have come across so far.

To support this conclusion I offer a general observation about econometric analysis with observational data, and then some specific criticisms of the data used in the automobile literature. The general observation is that economists sometimes tend to be unduly confident about the reliability of conclusions derived from econometric analysis using incomplete, observational data. The rebound effect is an example of what would be considered a treatment effect, and the ideal way to measure this would be through a controlled experiment. Since this is not possible, we are forced to rely on observational data, but these are always likely to be

susceptible to frailties stemming from basic uncertainty regarding the true relationship and a slew of data problems including collinearity among the variables that are measured, error in the measurement of these variables, the omission of other, relevant variables, and dependence between the some of the variables omitted and some of those included. Freedman (1991, 1999) presents a powerful exposition of the problems these phenomena create for the reliable detection of causation.

These phenomena certainly apply to the econometric literature on the rebound effect. I will mention several specific examples. With regard to uncertainty about the true behavioral relationship, I have already mentioned the fact that most studies impose the assumption that vehicle utilization responds identically to changes in the price of gasoline and in mileage per gallon, without testing whether this is true. Another example is the possibility that the behavioral response is asymmetric with respect to an increase versus a decrease in gasoline price: many studies (including Small and Van Dender) impose this assumption without testing whether it is true. A third example is that the literature generally assumes that a single behavioral relationship applies to the data being modeled – apart from the few shift variables that are included in the regression, all drivers in the state, or in the nation, respond in the same way to a change in fuel cost. To the extent that there are separate behavioral relationships for distinct sub-populations, their relative share of the overall population may change endogenously because of the differences in their response behavior, thereby changing the relationship observed in the aggregate and causing aggregate regression coefficients to be unstable.

With regard to data problems, one example is that the price variable in the demand for vehicle utilization is measured incorrectly. The existing studies treat the cost of driving as though it consisted exclusively of the cost of gasoline, but this is only a portion of the variable cost of operating a vehicle. As Gerard and Lave (2003) point out, using the Ford Taurus as an example, according to the American Automobile Association the costs of driving a Taurus include 7.1 cents per mile for gasoline and oil (not just gasoline), 4.1 cents per mile for repair and maintenance, and 1.8 cents per mile for tire wear. The pure gasoline component – which is what is modeled in the econometric studies – accounts for about 7 cents out a total of 13 cents. In addition, for each mile driven beyond the first 15,000 miles, AAA estimates that the Taurus depreciates at a rate of 18.8 cents per mile. The pure gasoline component is only about 22% of the total variable cost of 31.8 cents per mile. Moreover, this fraction is likely to vary for different vehicles and for different types of driving conditions (urban versus rural, highway versus city street). Hence, using gasoline price as the measure of the cost of driving introduces a substantial random error, but one with a distinctly non-zero mean. Another problem with most studies of the demand for vehicle utilization, Goldberg (1995, 1998) being the notable exception, is the omission of vehicle attributes that are highly likely to influence this demand. Many of the attributes that are believed to influence model choice in the vehicle purchase decision also have the potential to influence vehicle utilization – if the car feels safe, is comfortable, can accommodate all the children, etc these are considerations that could influence utilization decisions (e.g., whether to drive or fly when visiting grandmother) every bit as much as the cost of gasoline. These variables are omitted because of the lack of data, but that does not render their omission harmless. My understanding is that there have been some distinct trends in such vehicle attributes over the past 30 years, and these may perchance covary with some of the trends in gasoline prices over this period. Consequently, the omission of vehicle attributes from

the driving demand model is likely to be a substantial source of bias when estimating the rebound effect associated with fuel efficiency.

For these reasons, I am skeptical of the reliability of most of the existing estimates of the rebound effect associated with fuel efficiency. Nevertheless, as I indicated above, it seems likely to me that that the rebound effect associated with the climate change emission reduction regulations is likely to be small if not zero.

Response: We divide this comment into its three main observations, restated below in our words.

1. *All econometric studies are subject to a variety of data problems including omitted variables that are correlated with observed variables and therefore may cause measured coefficients to be biased.*

Response: This observation is true, but we would not be quite as pessimistic about the possibility of learning anything from economic data. These are problems inherent in *all* econometric studies—in fact, in almost every kind of empirical investigation, although it may not always be as apparent because assumptions are not always as explicit as they are in econometric studies. Nevertheless, applied economics has been very successful in shedding light on economic policy questions. Partly this is because even when specific findings need modification due to one or more of the problems outlined, they reveal logical relationships and usually at least a valid first approximation to the magnitudes of causal effects. In the case of studies of the rebound effect, there seems a consistent tendency to converge toward a rather small number as studies become more sophisticated and use more data, even though they differ widely in their exact specification. This tendency makes it likely that the results are not far off the mark.

2. *Studies of automobile usage in particular have a set of specific examples of these problems:*

   (a) *Assumption that vehicle use responds identically to fuel price and to fuel efficiency;*

   (b) *Possibility that response over time would be asymmetric with respect to increases or decreases in driving cost;*

   (c) *Omission of components of variable cost other than fuel;*

   (d) *Omission of vehicle attributes that may covary with fuel prices.*

Response: We can take up each alleged weakness in turn:

(a) We agree, as stated in the report on p. 45, that our evidence does not confirm the hypothesis that people respond identically to fuel price and to fuel efficiency. Given the nature of the evidence, what it shows is that people's response to fuel prices can be measured fairly precisely and appears to be quite small; but that the response to fuel efficiency may be even smaller and, as we stated and Hanneman reiterates, cannot be proven to be different from zero. These are valuable observations for policy. We think policy toward greenhouse gas emissions is most

sensibly designed by taking the more cautious view that the theoretical assumption is valid, while recognizing that the true situation may be more favorable for the effectiveness of the regulations than under this assumption.

(b) Although there may be asymmetric responses to price increases and price decreases in the short run, these are unlikely to persist as the vehicle fleet turns over. Therefore we do not think asymmetry would make much difference over a period of a decade or so. Furthermore, we believe that attempts to measure asymmetry in the literature have not been very successful because it has been quite difficult to specify an appropriate dynamic model. This is now stated on p. 33.

(3) The price variable used in these studies is, quite appropriately, fuel cost, not total variable cost. Thus it is not "measured incorrectly." Rather, the omission of oil, tires, maintenance, variable depreciation, and other variable costs is just another example of point (1). There is no reason to think that this omission causes any more problems than any other omitted variable, for example one measuring the precise nature of transit service available. Even if we included data on such costs, they would not necessarily affect usage identically to fuel costs because of difference in timing and visibility, so we would not want to simply add them to fuel cost. Such an additive structure would reflect a theoretical assumption of the same nature as the one concerning the identical effects of fuel efficiency and fuel price. We think that omitted time costs are more important than any omitted money costs, and these are likely to grow with income and congestion. This is why we included measures of income and urbanization and interacted them with the fuel-cost elasticity of usage, so that some of the effects of omitting them could be captured in that way. This procedure was successful in identifying a relationship between the cost elasticity (i.e. rebound effect) and income, which in turn plays an important role in our projections for California between 2008 and 2020.

(4) We did construct variables to measure general trends in vehicle quality, consisting of measures of the diffusion of several specific technologies. However none of these were found to have a measurable effect when we also include a simple linear time trend. We therefore reported results with just the time trend, which we believe captures most of the trends in vehicle attributes, as now noted on p. 33. Because fuel prices move over time in a quite variable way, and also vary across states, we think it is unlikely that unmeasured vehicle attributes are closely correlated with fuel prices or with fuel cost per mile.

3. *Professor Hanneman believes it is nevertheless likely that the rebound effect is small, or even zero, consistent with our results.*

Response: We agree that the rebound effect is most likely small (although on theoretical grounds we doubt it is zero). We are pleased that we have added some additional evidence to identify this behavioral parameter, and that we have for the first time provided evidence on how it might vary with income over space and time.