Historical Evidence for Energy Consumption Rebound in 30 US Sectors and a Toolkit for Rebound Analysts

Harry D. Saunders

List of Data Caveats and Assumptions

estimate of the actual. To provide a fair comparison among rebound trajectories, the same model was used to project actual energy consumption as was used to project the 100% rebound and zero rebound trajectories. In individual sectors, projected actual energy consumption deviates slightly from the true actual, sometimes somewhat above, sometimes somewhat below. However, in aggregate the estimated actual is virtually indistinguishable from the true actual, as shown in Figure 5.

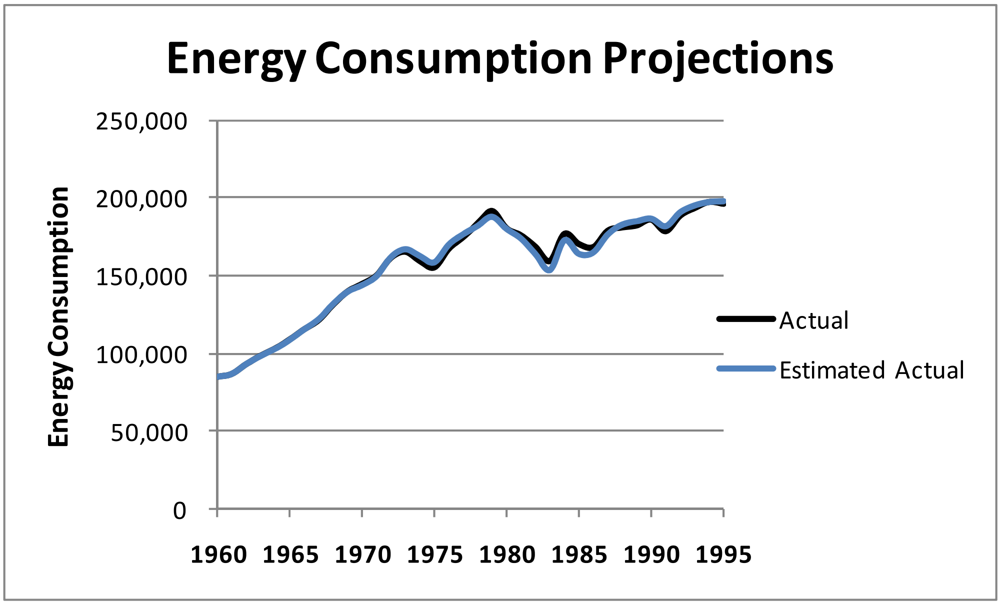

Figure 5. Actual vs. Estimated Actual Energy Consumption Trajectories

Given the multitudinous methodological elements involved in modeling these energy consumption trajectories, and the multiple excursions required back and forth between the primal and dual worlds, this is perhaps a testament to the reliability of the methodology, and perhaps a statement about the basic validity of the model's underlying assumptions of profit maximization and perfectly competitive firms.

Cautions and Limitations

The asserted strengths of the methodology do not mean it is without significant limitations. To enable econometrically-measured rebound estimates requires, at this point in methodology development, certain compromises. In addition to relying on the standard assumptions of microeconomics such as profit maximizing producers, perfect competition, and continuity, homogeneity and monotonicity of production/cost functions, the approach is limited in the following ways:

- The chosen functional form of the cost function is not the best possible candidate for an analysis of this type. While it is clearly superior to the CES (Solow) function and is highly popular among analysts, the Translog function is not as general as other forms developed by economists. As discussed elsewhere (Saunders, 2008), an ideal candidate would be the Gallant (Fourier) form, since this function does not presume membership in

- A limitation imposed by the use of Translog forms is that forcing concavity (a requirement generally called for by economists) presents some difficulties. It has been shown elsewhere (Saunders, 2008) that forcing global concavity on a Translog function necessarily leads to energy consumption backfire.[16] This limitation can be overcome by introducing constraints on the econometric measurement (discussed in the Methodology section) to force local concavity only. Ryan and Wales (2000) have provided such a methodology. However, this requires that the measured cost function be tested to see how well it behaves, concavity-wise. With the Ryan and Wales method, the measured cost function must be tested year-by-year post-measurement for each sector to evaluate its concavity over the relevant time horizon (and it depends on the reference year specified, since a different choice of reference year produces a different measured cost function). A rule-of-thumb metric is used to evaluate conformance with concavity.[17] But the results reported here include cases where not all sectors' cost functions honor concavity for every year.
- The factor technology gains measured in the analysis are assumed to be uniform over time, with smooth percentage technology gains for each factor each year. In reality, technology gains are more likely "lumpy," with new technologies appearing periodically in each sector. While the methodology can in principle accommodate year-by-year changes in technology, the added econometric parameters required would seriously degrade the statistical performance, and the statistical metrics generated using uniform technology parameters are already at the edge of respectability.
- For purposes of aggregating across sectors, a particularly simple assumption is made to permit conformance with general equilibrium theory. That is, consumer utilities are assumed to be of the Cobb-Douglas form. This creates a delightfully easy way to aggregate across sectors. With Cobb-Douglas utilities, when output prices and quantities change due to technology gains, demand for sectoral outputs can be treated independently and then summed together.[18] This provides a way to aggregate sectoral results that is consistent with general equilibrium conditions—but it must be skeptically considered a "poor man's" general equilibrium model.[19] A number of researchers (e.g., Turner, 2008,

---

[16] See Translog backfire theorem, Appendix F of Saunders (2008).

[17] Specifically, this metric is the sum of the positive eigenvalues of the core Hessian of the measured cost function over the relevant time horizon (modified as shown in Appendix A to comprehend the Ryan and Wales local constraints)—the larger this metric, the greater the departure from concavity.

[18] See, for example, Luenberger (1995), p. 132. Sector demand depends only on the sector-specific price of output.

[19] Robert Solow has called the Cobb-Douglas function the "Santa Claus" function. This application of it is yet another example of the "gifts" it provides analysts.

Wei, 2007,2010, Barker et al., 2007a,2007b) are hard at work to create more sophisticated, and realistic, general equilibrium models of rebound. So the aggregation presented here is best treated with considerable caution. That said, theoretical indications are that rebound is augmented to the extent substitution elasticities among factors is larger.[20] With the possibility of added factor substitution *among* sectors arising in a more robust general equilibrium model, this suggests that the Cobb-Douglas assumption may lead to an *understatement* of rebound magnitudes. Against this, because the Cobb-Douglas assumption implies unitary elasticity of demand for output, it could lead to an *overstatement* of rebound.

- In a similar vein, the analysis makes factor aggregation assumptions that are less than ideal. That is, while the cost of output, capital cost, factor demands and output levels are treated endogenously, labor and materials supply are treated as perfectly elastic. More specifically, prices applied in the rebound model for labor and materials are held fixed in nominal terms as between actual, 100% rebound, and zero rebound cases. This is not especially problematic for treatment of sectors individually since an individual sector is unlikely to have significant impact on factor supplies, especially those traded globally, and it is common to treat individual sectors as "price-takers." But aggregating across multiple sectors could be seen as more problematic. Since output levels will in reality generally be higher in the actual case than the 100% rebound case (due to factor efficiency gains), pressures on labor and materials supply will be greater, leading to higher labor and materials prices. However, this would have the effect of reducing the relative energy price, thus leading to a higher trajectory of energy consumption in the actual case, thereby *increasing* energy consumption rebound. It was felt that incorporating labor and materials supply functions in the analysis would run the risk of introducing an arbitrary element that could distort energy rebound in inscrutable ways. The methodology can accommodate depictions of labor and materials supply functions, but the risk of introducing what could be seen as arbitrary, or at least highly disputable, assumptions seemed high.[21] The bottom line: treating labor and materials supply as perfectly elastic likely *understates* actual energy consumption rebound.
- For energy, the situation is a little more complicated. A 100% rebound trajectory is one in which energy use is (generally) higher than in the actual case, at least for energy-specific technology gains. If a positive energy supply elasticity were introduced (undoubtedly the case in reality), the energy use trajectory in the 100% rebound case would accordingly be lower than if energy price remains unchanged, thus *increasing* rebound magnitudes. Offsetting this, a zero rebound condition would correspond with reduced energy consumption, thus reducing energy price and creating a higher trajectory of energy consumption in zero rebound conditions than without consideration of energy

---

[20] Saunders (1992,2008).

[21] To bound this problem, an attempt was made to include a feature in the Rebound Measurement Module that introduces perfectly *inelastic* supply for both labor and materials. Unfortunately, this is too severe, and invariably leads to instabilities in the projections.

price effects, and would thus *reduce* rebound magnitudes. This would have the effect of *reducing* the distance between the zero rebound and 100% rebound trajectories. While these effects may or may not be offsetting, for this analysis it was deemed prudent, as with labor and materials, to remain agnostic on the dynamics of energy supply (especially given the presence of OPEC as a non-competitive energy producer that sets global energy prices, making energy supply elasticity a complex concept), leaving for future analysts the task of comprehending the relative magnitudes of these offsetting effects. The toolkit methodology can accommodate an energy supply function, but the objective was to avoid hidden effects based on assumptions open to high controversy and dispute. Nonetheless, it is possible to argue that the rebound magnitudes reported in this article are distorted by this assumption (although whether the distortion understates or overstates rebound is a question that remains to be decided). That said, the assumptions of perfect (nominal) elasticity of labor, materials, and energy supply provide the means to aggregate across sectors in a way consistent with general equilibrium theory. Future analysts will no doubt have much to contribute here.

- The output of every sector is assumed to be directly consumed by end users. In reality, sectors use outputs from other sectors as inputs to produce their outputs. This “nesting” of outputs and inputs is entirely ignored in this analysis. While it is unclear what effect this shortcoming has on the magnitudes of measured rebounds, theoretical considerations suggest that it leads to an *understatement* of rebound. In particular, Lowe (2003) has shown that energy substitution elasticities become larger the more levels of nesting occur. As noted previously, theory indicates that larger energy substitution elasticities are associated with larger rebound magnitudes. Researchers such as Turner (2008,2009), Anson and Turner (2009), Allan et al. (2006), Hanley et al. (2006), Grepperud and Rassmussen (2004) and others are currently using models that better comprehend this “nesting” phenomenon. Lecca et al. (forthcoming) directly explore the consequences of different nesting schemes, helping fill a major gap in the field.
- The capital vintaging approach uses a “putty-clay” model. That is, while the newest vintage is deemed entirely flexible in choosing among the production possibilities represented by the cost/production function, once it is in place it is assumed to exhibit fixed factor and output capacity proportions, although factor and output magnitudes decline over time owing to depreciation. This is tantamount to considering older vintages as exhibiting Leontief technology behavior. This overlooks the potential for capital in place to be retrofitted in a way that changes its factor proportions. However, to the extent this reflects reality, it also suggests the potential for increased factor substitution across capital vintages. As noted above, increased factor substitution potential is theoretically associated with increased rebound, so the “putty-clay” approach likely *understates* rebound. This “putty-clay” approach is also undoubtedly more accurate than assuming new capital and capital in place are entirely fungible in a way governed by some sector-wide production function that includes both new and old vintages.

- Not all sectors represented in the Jorgenson et al. database are included in the analysis. Specifically, the analysis excludes coal mining and oil and gas extraction. The Jorgenson et al. data sets treat coal, oil, and gas as inputs to production in these sectors. While this makes perfect sense from a value-added perspective, these energy inputs are not actually consumed in these sectors, and so do not contribute to greenhouse gas emissions.[22] Similarly, the energy conversion sectors, petroleum and coal products (largely oil refining) do not consume the energy input, but rather transform it, so these sectors have been excluded as not consuming energy and creating associated emissions. The trade sector has also been excluded. The rationale for excluding it arises mostly out of the work of Allan et al. (2006) who have shown that this sector exhibits somewhat quirky behavior as regards rebound phenomena. That is, they have shown that energy efficiency gains in products that are exported can lead to greatly exaggerated local rebound effects. Government enterprises have been included even though it is doubtful this sector adheres to the assumption of profit maximization and perfect competition.
- The analysis considers only so-called “direct” rebound effects. That is, it includes both the output/income components of rebound and the substitution/intensity effects, both of which arise in the productive part of the economy. And it implicitly comprehends the phenomenon of consumers using savings from energy to purchase the output of other sectors (“indirect effects”).[23] But so-called “macroeconomic” effects are excluded. The term “macroeconomic” is fraught with confusion and conflicting uses, but one clear example of a “macroeconomic” effect is that such as might arise when energy efficiency gains provide the basis for as yet unforeseen new energy-using applications, products, enterprises or even whole new industries. (This should probably be called the “Jevons effect.”[24]) Accordingly, it is possible to argue that the results reported here thereby *underestimate* the economy-wide rebound arising from energy efficiency gains.
- There is judgment involved. The analysis relies on choosing among multiple theoretically plausible methods for depicting capital formation, new output-augmenting capacity, utilization rate profiles, and choice of reference year. For each sector, approximately 100 different method combinations are tested against the metrics of statistical performance, adherence to concavity, and minimum deviation of forecast projections of factor uses and output from actuals. That said, a remarkable thing is that method combinations that deliver good performance against one metric (e.g., concavity) also tend to deliver good performance against all metrics. Further, where performance is

---

[22] Energy is actually consumed in these sectors, and emissions produced, but it is not possible to separate out from the data set the portion of energy inputs actually consumed.

[23] Although with the Cobb-Douglas utility specification on consumer behavior assumed here, consumption is not reallocated among sectors.

[24] A good example of the “Jevons effect”: Tsao et al. (2010) have shown that new applications in efficient lighting have, since the 1700s, offset the energy efficiency gains from new lighting technologies almost exactly, leaving energy intensity of lighting unchanged over hundreds of years and independent of “luminous efficacy.” New lighting applications have continually arisen that offset energy consumption reductions due to energy efficiency gains, for more than 300 years, across 3 continents and across 6 technologies.

good for more than one method combination, measured rebound magnitudes are relatively stable and do not differ much one from the other. Accordingly, while choosing among method combinations involves judgment, it is unlikely that other researchers using this methodology would choose significantly different method combinations or report significantly different rebound magnitudes. Moreover, where more than one methodology provides reasonable explanatory power and plausible statistics, the methodology resulting in *lower* rebound has been used. Thus, the results may for this reason *understate* rebound magnitudes.

More generally, extreme effort has been made to avoid "cheating" whenever judgment is brought into play.[25]

- No consideration is given to producers employing "rational expectations" in their decision making. Rather producers are assumed to choose production technologies based on factor prices prevailing in the year the investment is made. While this is consistent with most models of energy consumption, it is a limitation.
- The analysis excludes consideration of what would have happened were carbon taxes or additional energy use regulations invoked during this time period. These have clear implications for forecasting future energy consumption trends and rebound effects. However, it is to be noted that such government interventions have the certain effect of reducing economic welfare (at least welfare narrowly construed to exclude externalities). Instead, this analysis indicates the effects on energy consumption of technology gains that do not come at a cost to economic activity. This seems the most honest way to evaluate pure rebound effects.
- Government monetary and fiscal policy is held fixed across the three rebound scenarios. In reality, for instance, had the 100% rebound case actually obtained, government may have invoked monetary or fiscal stimulus to offset the lower output trajectory associated with this case. Again, however, ignoring this seems the most honest way to provide an "apples-to-apples" comparison that isolates rebound effects.
- A final, important, caution: these results should **not** be taken as an argument against deploying new energy efficiency technologies. Such technologies increase economic welfare (narrowly construed to exclude externalities). It is just that they may not deliver the reductions in energy consumption presumed by many.

This impressive list of limitations should not be dismissed out of hand by rebound analysts as minor or irrelevant. Rather, the intent in delineating them is that practitioners need to find ways to overcome them if energy consumption rebound is to be properly understood.

---

[25] Robert Solow in the late 1970s gave a presentation at Stanford University called something like, "How to lie with econometrics" in which he showed the many ways in which it is possible for econometricians to fool themselves—or others. This had a profound and lasting effect on a young researcher.