Journal of Agricultural and Applied Economics, 36,3(December 2004):549–558

Do Economic Restrictions Improve Forecasts?

Elizabeth Murphy, Bailey Norwood, and Michael Wohlgenant

A previous study showed that imposing economic restrictions improves the forecasting ability of food demand systems, thus warranting their use even when they are rejected in-sample. This article evaluates whether this result is due to economic restrictions enhancing degrees of freedom or containing nonsample information. Results indicate that restrictions improve forecasting ability even when they are not derived from economic theory, but theoretical restrictions forecast best.

Key Words: demand systems, economic restrictions, forecasting, representative consumer

JEL Classifications: B4, C1, C3, C5

Using several popular demand systems in conjunction with food consumption data, Kastens and Brester (KB) have shown that theory-constrained demand systems forecast better out of sample (hereafter, "forecast") than their unrestrained counterparts. Although at first this seems to provide some justification for imposing theoretical constraints, it does not address the question of whether the forecast benefit derives from economic theory or from higher degrees of freedom.

Parameter restrictions serve to enhance degrees of freedom, regardless of whether the restrictions are derived from theory. Because models with greater degrees of freedom often forecast better, in this article we explore whether the theory-constrained models of KB forecast better because the restrictions are "true" or because their degrees of freedom are higher.[1] We wish to separate the contribution of forecast improvements due to economic theory from that of higher degrees of freedom.

We use the data from the KB study to reestimate their models with arbitrary restrictions. These arbitrary restrictions are not derived from theory, but they increase the degrees of freedom by an identical amount as the economic restrictions. Results indicate that arbitrary restrictions, because of more degrees of freedom, do improve forecasts relative to no restrictions. However, economic restrictions improve forecasts even more, which suggests that there is valuable information con-

Elizabeth Murphy is research assistant, Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, NC. Bailey Norwood is assistant professor, Department of Agricultural Economics, Oklahoma State University, Stillwater, OK. Michael Wohlgenant is William Neal Reynolds Distinguished Professor, Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, NC.

The authors thank Terry Kastens for providing the data, computer code, and notes necessary to replicate his results and Nicholas Piggott for helpful comments. Comments and insights by two anonymous reviewers are gratefully acknowledged as well.

[1] Suppose Models A and B are approximations to a true unknown functional form. Sawa shows that, although Model B may be a closer approximation in a large sample sense or its structure may resemble the true form more, in small samples, model A may be a better approximation if its degrees of freedom are larger. Appeals to model selection criteria that contain penalty parameters are typically made based on this fact.

tained in economic theory and that economic theory has an important role in forecasting.

The Value of Parameter Restrictions

It may seem strange that theoretical restrictions would be rejected in sample and then reduce forecast errors out of sample.[2] Why would theoretical restrictions appear to be informative out of sample but not in sample?[3] One reason, based on the concept of sample and nonsample information, is that economic restrictions improve forecasts because economic theory is informative. The other explanation, based on degrees of freedom issues, is that any restriction might improve forecasts, regardless of whether the restriction is true or not.

Sample information refers to a set of observations. Theoretical restrictions are a form of nonsample information. They represent information that researchers believe to be true but may not be reflected adequately in a random sample. The three most popular restrictions—symmetry, homogeneity, and adding up—are derived from the theory of the representative consumer. Their derivation rests on several assumptions that may be too restrictive. Assumptions commonly made are that all consumers possess and maximize the same utility function, the parameters of that function are invariant with time, all consumers face identical real prices, and either all specified goods must be exhaustive or a subset must be separable (Deaton and Mullbauer).

But restrictions derived from economic theory need not hold perfectly to have value. Economic restrictions convey information even if none of the above assumptions hold. If beef is a strong substitute for pork, pork should be a strong substitute for beef. The symmetry condition ensures that this is the case.[4] The point is that theoretical restrictions may convey much information that we know about consumers, even if their parametric representation is not perfectly accurate.

Suppose we wish to estimate a parameter vector β for use in forecasting, and the prediction errors (either in or out of sample) are an increasing function of the distance between the true vector β and its estimate, $\hat{\beta}$. The more information contained in $\hat{\beta}$, the smaller this distance. Information in $\hat{\beta}$ is a function of sample and nonsample information. Let the unrestricted estimate be denoted $\hat{\beta}_U$ and its theory-constrained counterpart be $\hat{\beta}_R$, where $\hat{\beta}_U$ only contains sample information and $\hat{\beta}_R$ contains sample and nonsample input. When predicting in-sample observations of the dependent variable, it is possible that sample information may dominate the nonsample information (the information in theoretical restrictions). The exact formulation of a demand system is unlikely to be correct. Imperfect competition, product differentiation, and consumer heterogeneity prohibit us from identifying the exact demand function. Even if an economic restriction is true, when combined with an imperfect demand function, those restrictions may be rejected. In these cases, allowing the estimation routine to search unrestricted over all possible values for β results in significantly smaller (in-sample) prediction errors than if constrained by theory.

Now, let us turn to the case where $\hat{\beta}_U$ and $\hat{\beta}_R$ are used for forecasting out of sample. Specifically, we focus on the case where observations from earlier dates are used to forecast future observations. It is likely that the true parameter vector β changes over time because of changing consumer preferences, model misspecification, and other complexities involved

---

[2] "In sample" refers to the set of observations used to calculate parameter estimates. For instance, the likelihood ratio test is an in-sample statistic, because the likelihood functions are calculated using the same observations in which the parameters were generated. Conversely, "out of sample" refers to a set of observations that were not used to estimate parameters. For instance, if data from 1970–1990 are used to predict prices in 1970–2000, the predictions for 1970–1990 are in sample and those for 1991–2000 are out of sample.

[3] In sample, "informative" is defined as not being rejected. If they are not rejected, then they may (but do not have to) be interpreted as true. Out of sample, something is "informative" if it reduces forecast errors.

[4] That is, assuming that the budget shares of both goods are small, eliminating effects from income elasticities.

Table 1. Forecasting Performance of an AIDS Model under Alternative Parameter Restrictions

Forecast Horizon	Food Group					
	Meats	Eggs	Dairy	Fats	Cereals	Sweets
	Ratio of root-mean-squared forecast errors of unrestricted to theoretically constrained AIDS demand systems[a]					
1 Year[b]	1.25	0.96	1.24	0.98	1.37	1.23
1–11 Years[c]	2.09	1.29	1.29	1.33	1.32	1.45
	Ratio of root-mean-squared forecast errors of unrestricted AIDS demand systems to parsimonious AIDS systems[d]					
1 Year[b]	0.72	1.17	0.87	1.07	0.91	0.93
1–11 Years[c]	0.68	1.10	1.27	1.09	1.53	0.95

Notes: See Kastens and Brester for more details on the data and model specification. All forecasts were performed identical to the method of Kastens and Brester.

[a] This model is denoted by FDLA/ALIDS in Kastens and Brester. The theoretically constrained system imposes symmetry and homogeneity, both of which were rejected using likelihood ratio tests.

[b] This means the quantity of the food group in year t was forecasted using observations from the previous 25 years.

[c] This means the quantity of the food group in years t through $t + 10$ were forecasted using observations from the 25 years previous to year t.

[d] The parsimonious AIDS model sets all parameters except the own-price and intercept terms to zero. This model was not used in Kastens and Brester.

in econometrics, in ways that are difficult to capture even with the most advanced random-coefficient estimation techniques. If this is true, then sample information from previous time periods are of less use in explaining future observations than they were in explaining in-sample observations. But the value of nonsample information via theoretical restrictions stays the same, because theory is not time dependent. The amount of information in theoretical restrictions, relative to the information contained in the in-sample observations, is now greater, and the restricted estimates' forecasting ability, relative to unrestricted estimates, begins to improve.

Some evidence for this is given in Table 1, using data from KB and their form of the AIDS model. This table shows the ratio of forecast errors from an unrestricted AIDS model to an AIDS model with symmetry and homogeneity imposed. With only a 1-year-ahead forecast horizon, the restricted model performed better in some cases and worse in others. Once this horizon increases, the restricted form has lower errors for all food groups. As the forecast horizon increases, the theoretically constrained model forecasts better. This may be due to the economic content of the restrictions, i.e., that the restrictions are theory based and the theory is sound.

Restrictions do not have to be based on theory or empirical results or even make sense to improve forecasts. Restrictions may improve forecasts simply because they increase the degrees of freedom (Brieman). As Sawa notes, even if one model is a closer approximation to the true model analytically, in small samples, models with more degrees of freedom may better represent the true data-generating process. Consider again the data and AIDS model used by KB. In Table 1, an unrestricted AIDS model is compared with a parsimonious AIDS model, where the value of all parameters except for own price and intercept terms are set to zero. At a 1-year horizon, the unrestrained AIDS model has lower forecast errors for four of six goods, but at a 1–11-year horizon, the parsimonious AIDS model has better forecasts for four of six goods. At longer forecast horizons, forecast improvements can be obtained simply by increasing the degrees of freedom. This finding is not isolated; it is generally accepted that models with more degrees of freedom tend to forecast better.

Consider again one forecast series from the parameter vector $\hat{\beta}_U$ and one from the vector

$\hat{\beta}_R$. In this case, it is assumed that $\hat{\beta}_R$ is estimated using restrictions not based on theory, but because restrictions are imposed, the degrees of freedom are higher for $\hat{\beta}_R$ than for $\hat{\beta}_U$. The mean-squared error of $\hat{\beta}_R$ from its true value β is the variance of the estimator plus the squared bias, i.e., $E(\hat{\beta}_R - \beta)^2 = V(\hat{\beta}_R) + [E(\hat{\beta}_R) - \beta]^2 = V_R + \text{BIAS}_R^2$. Forecasts from $\hat{\beta}_R$ will be more accurate than those from $\hat{\beta}_U$ if $V_R + \text{BIAS}_R^2 < V_U + \text{BIAS}_U^2$. If the restrictions are not true, it is certainly the case that $\text{BIAS}_R^2 > \text{BIAS}_U^2$. However, because degrees of freedom are higher for the restricted estimate, it may be that $V_R < V_U$, such that the mean-squared error for $\hat{\beta}_R$ is lower than that for $\hat{\beta}_U$, thus producing better forecasts. A case can be made for V_R being lower than V_U. With more degrees of freedom, the restricted parameter estimates are derived from more observations; thus, their variability in repeated samples should be smaller (Breiman).

The KB study found that models with economic restrictions forecast better than their unrestrained counterparts. We have just explained how this could occur. First, the economic theory used to derive those restrictions might be valuable nonsample information, i.e., the theory might be correct. Second, even if the theory is not correct, restrictions serve to increase degrees of freedom, and more degrees of freedom could result in more accurate forecasts. Which explanation is correct? This is an important question to address, because the answer will guide economists as to whether improved forecasts can be obtained by developing better theories, using more parsimonious models, or both.

A simple method can be used to address this method. This method entails estimating demand while imposing arbitrary restrictions that have no reason to be true but increasing degrees of freedom by an equal amount as economic restrictions and then comparing those forecasts to a scenario where economic restrictions are imposed. If theoretical restrictions provide better forecasts than these arbitrary restrictions, then we can say the theoretical restrictions contain nonsample information useful for forecasting. If they do not, we must conclude that all forecasting improvements in KB are due to increases in degrees of freedom and not theory. We performed this test using the exact data and estimation methods of the KB study. These methods are discussed in the next two sections.

Demand Systems with Economic and Arbitrary Restrictions

The Rotterdam, AIDS, and the first-difference-double-log (FDDL) models are demand systems used by KB for food-demand analysis. Six food groups were used: meats, eggs, dairy, fats, cereals, and sweets. A seventh group, "all other goods," was also constructed. Thus, there are $i = 1, \ldots, 7$ exhaustive goods, where the price and quantity of those goods are denoted by p_i and q_i, respectively. Denoting per-capita nominal income by X, the Rotterdam model is given by

$$(1)\quad \hat{w}_{i,t}\Delta\ln(q_{i,t}) = \alpha_i + \sum_{j=1}^{N}\gamma_{i,j}\Delta\ln(p_{j,t}) + \beta_i\left[\Delta\ln(X_t) - \sum_{j=1}^{N}\hat{w}_{j,t-1}\Delta\ln(p_{j,t})\right] + \varepsilon_{i,t},$$

where $\hat{w}_{i,t}$ is the average expenditure share,

$$(2)\quad \hat{w}_{i,t} = \frac{1}{2}w_{i,t} + \frac{1}{2}w_{i,t-1} = \frac{1}{2}\frac{p_{i,t}q_{i,t}}{X_t} + \frac{1}{2}\frac{p_{i,t-1}q_{i,t-1}}{X_{t-1}}$$

and Δ is the across-period difference operator. The version of the AIDS model used by KB is

$$(3)\quad \Delta w_{i,t} = \alpha_i + \sum_{j=1}^{N}\gamma_{i,j}\Delta\ln(p_{j,t}) + \beta_i\left[\Delta\ln(X_t) - \sum_{j=1}^{N}w_{j,t-1}\Delta\ln(p_{j,t})\right] + \varepsilon_{i,t}.$$

The FDDL model is

$$(4)\quad \Delta\ln(q_{i,t}) = \alpha_i + \sum_{j=1}^{N}\gamma_{i,j}\Delta\ln(p_{j,t}) + \beta_i\Delta\ln(X_t) + \varepsilon_{i,t}.$$

Each of the N goods has one equation, and each equation has $N + 2$ parameters. If estimated as a system, one equation must be dropped, leaving $N - 1$ goods.[5] The remaining system contains $(N - 1)(N + 2)$ parameters where each, if not constrained, must be estimated. The number of parameters to estimate can be reduced by imposing economic restrictions. Two examples are the homogeneity and symmetry conditions. The homogeneity and symmetry conditions decrease the number of parameters to estimate by $(N - 1)$ and $[(N - 1)(N - 2)]/2$, respectively. The adding-up condition is not imposed because it is automatically satisfied by the data.

To determine whether better forecasts using economic restrictions are the result of accurate theory, more degrees of freedom, or both, we compared models with economic restrictions to models with arbitrary (not derived from theory) restrictions. The arbitrary restrictions were chosen such that they enhanced degrees of freedom by the same amount as economic restrictions and in a similar way but were not based on theory and were randomly chosen.

For the AIDS and Rotterdam model, the symmetry condition states that $\gamma_{i,j} = \gamma_{j,i} \ \forall \ i \neq j$. The symmetry condition for the FDDL model can be stated as a linear function $\gamma_{i,j} = (w_j/w_i)\gamma_{j,i} - w_j(\beta_i - \beta_j) \ \forall \ i \neq j$. For generality, the symmetry restriction for all three models is written as

$$(5) \qquad \gamma_{i,j} + (a_{i,j})\gamma_{j,i} = b_{i,j} \ \forall \ i \neq j.$$

In the AIDS and Rotterdam models, $a_{i,j} = -1$ and $b_{i,j} = 0 \ \forall \ i, j$, and for the FDDL model $a_{i,j} = (-w_j/w_i)$ and $b_{i,j} = -w_j(\beta_i - \beta_j)$.

We created a set of arbitrary restrictions by replacing the symmetry restriction in Equation (5) with

$$(6) \qquad \gamma_{i,j} + (a_{i,j})\gamma_{k,r} = b_{i,j},$$

where both $i = r$ and $j = k$ cannot hold but one of them can. Because these restrictions can only be locally imposed on the FDDL model, they were imposed at the sample mean of budget shares.

Notice that Equations (5) and (6) appear similar and enhance degrees of freedom by an identical amount. The symmetry restriction states that if meat is a substitute for fat, then fat is a substitute for meat. This is a logical statement. An arbitrary restriction may state something illogical, such as "if meat is a substitute for fat, then eggs are a substitute for sweets." Although the arbitrary restriction contains no logic, it is a cross-equation parameter restriction with the same functional form as the symmetry restriction. For this reason, we believe that these arbitrary restrictions are as close as possible to the symmetry restriction while still containing no economic information.

There are numerous values of i, r, j, and k that would make Equation (6) a feasible arbitrary restriction. In fact, there are so many possible values that it is almost certain that one set of values would outperform the symmetry condition by chance alone. To prevent the results from being the result of a fortuitous choice of arbitrary restrictions, we compare forecasts from a model with the symmetry condition to the average forecast of that same model using 100 different arbitrary restrictions.

A computer program was written to randomly select values of (i, j) and (k, r) for the arbitrary condition. If both $i = r$ and $j = k$, then the randomly chosen values were discarded. Otherwise, the values were kept as an arbitrary restriction and new values were generated. The process continued until 100 unique arbitrary restrictions were constructed. For the remainder of this article, when we refer to a forecast using arbitrary restrictions, we are implicitly referring to the average forecast using the 100 different arbitrary restrictions.

Model Estimation and Forecasts

All data, estimation, and forecasts were performed identical to the method of KB, which

---

[5] If there are N exhaustive goods, then $\Sigma_{i=1}^{N} w_i = 1$. This results in the fact that, for the AIDS and Rotterdam models, if one tries to simultaneously estimate equations for all goods using conventional methods, the matrix of independent variables is singular, preventing unique parameter estimates. We elected to drop the "all other goods" good, and because its value is not directly observed, did not attempt to forecast it. Thus, although there are seven total goods in the demand systems, we only forecasted demand for six goods.

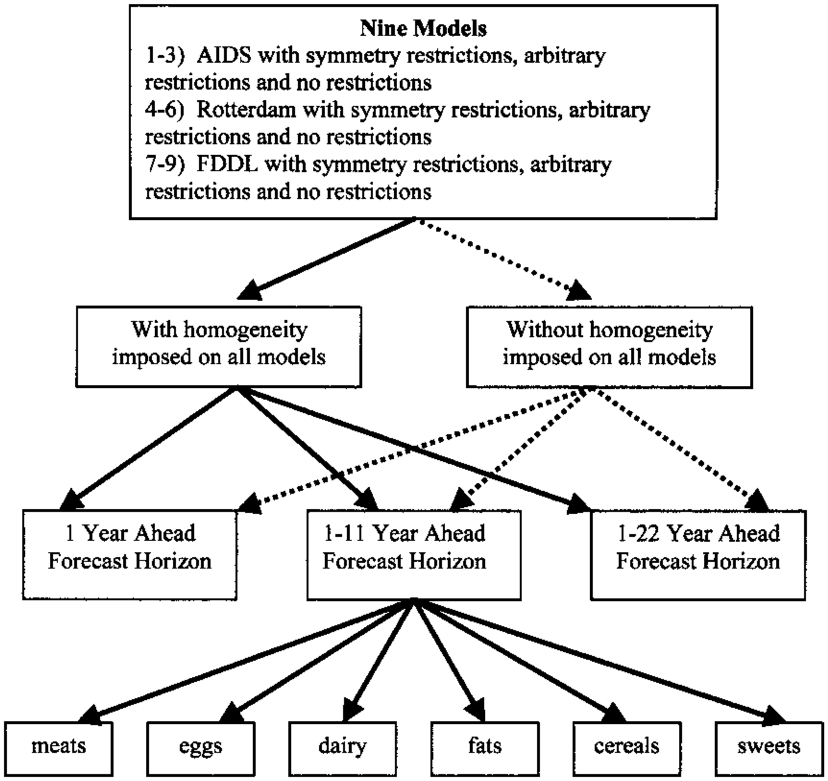

Figure 1. Models, Restrictions, Forecast Horizons, and Goods Used for Forecasting

are discussed thoroughly in their article. The original data and code were made available by the authors, and we compared our estimates with those published in KB, to ensure that our estimation procedure was the same. Thus, we defer most details to the KB article. Data covered the years 1924–1992, in which three model-updating methods were used. The first updating method used a 1-year horizon and a constant sample size of 25 observations. First, data from the years 1924–1948 were used for estimation and forecasting food quantities in 1949. Then, data from 1925–1949 were used to forecast in 1950. This continued until data from 1968–1991 were used to forecast in 1992. This provided 44 forecasts.

The second updating method used multiple horizons of 1 to 11 years. First, data from 1924–1948 were used to forecast in the years 1949–1959. Then, data from 1935–1959 were used to forecast in 1960–1970. Finally, the third forecast series horizon extended from 1 to 22 years, always maintaining a constant sample size of 25. Each of the three updating methods is referred to herein as a "forecast horizon." For each of the three forecast horizons and six food groups, there are 18 forecast series to compare for each model and restriction combination (e.g., 18 forecasts for the AIDS model with symmetry conditions, 18 forecasts for the AIDS model with arbitrary conditions, and 18 forecasts for the AIDS model with no restrictions). The combinations of models, restrictions, forecast horizons, and goods are illustrated in Figure 1.

The forecasting ability of models with symmetry and arbitrary restrictions were compared with and without homogeneity imposed on all models. Forecasts results were compared several ways. The first was identical to KB,[6] which ranked models according to which

[6] Specifically, we used the nominal ranking method as shown in table 5 of KB. Consider the first comparison as shown in Figure 1, where each three models are estimated with three different restriction sets (eco-

had the lowest mean-squared error across the three forecast horizons and six goods. The second method illustrated differences in the median-squared error between restriction types. To illustrate the median-squared error approach, let $(e_{M,G,S})^2$ be a squared forecast error from model M (M = AIDS, Rotterdam, or FDDL), good G (G = meat, eggs, etc.), and forecast series S (S = 1-year forecast horizon, 1–11-year horizon, or 1–22-year horizon). Consider a setting where we compare symmetry versus arbitrary restrictions with homogeneity always imposed. Let $(e^{S,H}_{M,G,S})^2$ denote the forecast error with symmetry and homogeneity and $E[(e^{A,H}_{M,G,S})^2]$ be the average forecast error across the 100 arbitrary restrictions (where homogeneity is imposed on all 100 models). A variable $P\{(e^{S,H}_{M,G,S})^2 < E(e^{A,H}_{M,G,S})^2]\}$ is constructed that equals one if true and zero if false. There will be a total of 1,908 unique values of $P(\cdot)$.[7]

The individual values of $P(\cdot)$ do not constitute a random sample. Their values are likely to be correlated across models, goods, and forecast horizons. However, P is still a good indicator of how well models with economic restrictions perform relative to arbitrary restrictions across all settings and so was used as guidance even though conventional statistical tests of $P(\cdot)$ would not be correct. It is especially useful because it provides one number for comparing forecasts across different goods and models. If the values of $P(\cdot)$ constituted a random sample, we could test the null hypothesis that $P = ½$, which states that the two types of restrictions forecast equally well. Specifically, it states that the median-squared forecast error using the two restriction types are identical.

If P is significantly greater than ½, the median-squared forecast error using economic restrictions is smaller. We would then conclude that economic restrictions do convey valuable information and have a greater use than simply enhancing degrees of freedom. The test is referred to as a nonparametric sign test, and the test statistic is $2T^{1/2}[P - ½]$ (Mendendhall, Wackerly, and Schaeffer), where T is the number of forecasts. This test is also used to evaluate whether the median forecast error is smaller using arbitrary restrictions or no restrictions. Again, this test is an indicator; the extent to which it can be considered valid is left to the reader.

Forecast Results

Table 2 shows the models that have the lowest mean-squared error across the 18 models, goods, and forecast horizons. A higher number corresponds to a higher ranking (lower mean-squared errors) across the three forecast horizons and six goods. As KB noted, FDDL is a better forecaster than the AIDS or Rotterdam models. Results indicated that models with restrictions forecast better than their unrestricted counterparts, regardless of whether those restrictions were derived from theory or arbitrary. These differences are statistically significant. Using AGS (Ashley, Granger and Schmalensee) tests, across all models, goods, and forecast horizons, the mean-squared error from unrestrained models was significantly higher than if arbitrary or economic restrictions are imposed 94% of the time.[8] This suggests that part of the reason economic restrictions improve forecasts is because they enhance degrees of freedom.

A more important result is that economic restrictions provided forecasts that were superior to arbitrary restrictions. Based on mean-

nomic restrictions, arbitrary restrictions, and no restrictions). This provides nine models to rank. The model with the lowest mean-squared error for a single good and forecast horizon is given a value of nine, and the model with the highest error is given a value of one. Because there are six goods and three forecast horizons, this implies a total of 18 rankings. Letting R_i (i = 1, . . . , 18) be the model ranking for a model and single good and forecast horizon, the overall nominal ranking for that model is $\Sigma_i R_i/18$. Thus, a higher value indicates a better model.

[7] Three models, six goods, three forecast series, and 44 forecasts per series implies 2,376 total forecasts. However, (26·3·6) 468 forecasts are redundant, leaving 1,908 unique forecasts.

[8] That is, for any good, model, and forecast horizon, models with no restrictions will have a significantly higher mean-squared error than models with arbitrary or economic restrictions in 94% of cases.

Table 2. Average Rankings of Demand Systems

Model	Restriction Type	Average Model Ranking
With homogeneity imposed on all models		
AIDS	Symmetry	4.11
AIDS	Arbitrary[a]	3.28
AIDS	No restriction	1.44
ROTTERDAM	Symmetry	5.78
ROTTERDAM	Arbitrary[a]	5.33
ROTTERDAM	No restriction	3.11
FDDL	Symmetry	8.56
FDDL	Arbitrary[a]	7.94
FDDL	No restriction	5.44
Without homogeneity imposed on any model		
AIDS	Symmetry	4.44
AIDS	Arbitrary[a]	3.17
AIDS	No restriction	1.39
ROTTERDAM	Symmetry	5.67
ROTTERDAM	Arbitrary[a]	4.72
ROTTERDAM	No restriction	3.39
FDDL	Symmetry	8.39
FDDL	Arbitrary[a]	7.67
FDDL	No restriction	6.17

Notes: The average model ranking is over 18 forecast contests (six food types times three forecast horizons). The highest ranked model (model with the lowest out-of-sample-root-mean-squared-error) in any ranking is assigned a value of nine, and the lowest ranked model is given a value of one. The reported rankings above are the average ranking for each model across all eighteen rankings. Thus, the higher the average ranking the better the model forecasts. A higher number indicates a higher average ranking/lower mean-squared forecast errors.

[a] Recall the models with arbitrary restrictions are a composite model of 100 individual models, each with a unique and randomly generated arbitrary condition.

squared error rankings shown in Table 2, economic restrictions out-forecasted arbitrary restrictions and no restrictions alike. AGS tests revealed that models with economic restrictions have significantly lower mean-squared errors than those with arbitrary conditions 67% of the time.

In Table 3, the forecast rankings are separated by the forecast horizon. Homogeneity was not imposed on any of these models. These results further confirm the previous findings but also reveal how forecasts using different restrictions perform as the forecast horizon lengthens. When comparing forecasts from the 1-year horizon with the 1–22-year horizon (see Table 4), one can see that the average ranking of models with economic restrictions always rises, whereas that with no restrictions always falls. This supports the hypothesis stated earlier that economic restrictions contribute more toward forecast accuracy the longer the horizon.

Table 3. Average Rankings of Demand Systems Across Forecast Horizons

Model	Restriction Type	Average Model Ranking
One Year Horizon		
AIDS	Symmetry	4.17
AIDS	Arbitrary	2.17
AIDS	No restriction	1.50
ROTTERDAM	Symmetry	5.70
ROTTERDAM	Arbitrary	4.70
ROTTERDAM	No restriction	3.50
FDDL	Symmetry	8.34
FDDL	Arbitrary	8.17
FDDL	No restriction	6.83
1–11-Year Horizon		
AIDS	Symmetry	4.67
AIDS	Arbitrary	3.67
AIDS	No restriction	1.33
ROTTERDAM	Symmetry	5.50
ROTTERDAM	Arbitrary	4.83
ROTTERDAM	No restriction	3.50
FDDL	Symmetry	8.33
FDDL	Arbitrary	7.33
FDDL	No restriction	5.83
1–22-Year Horizon		
AIDS	Symmetry	4.50
AIDS	Arbitrary	3.67
AIDS	No restriction	1.33
ROTTERDAM	Symmetry	5.83
ROTTERDAM	Arbitrary	4.67
ROTTERDAM	No restriction	3.17
FDDL	Symmetry	8.50
FDDL	Arbitrary	7.50
FDDL	No restriction	5.83

Notes: The ranking system is identical to that of Table 2, except that the average model ranking is over six forecast contests (six food types times one forecast horizon). Homogeneity was not imposed on any of these models. A higher number indicates a higher average ranking/lower mean-squared forecast errors.

Table 4. Tests for Significant Differences in Median-Squared Forecast Errors Using Economic, Arbitrary, and No Restrictions

Restriction Set A	Versus Restriction Set B	Forecast Horizon (Years)	Smaller Forecast Errors Using Restriction Set A Relative To Set B (*t*-statistic)
Symmetry and homogeneity	Arbitrary and homogeneity	1	64% (7.96)[a]
		1–11	65% (7.83)[b]
		1–22	65% (5.93)
Symmetry only	Arbitrary only	1	64% (7.88)
		1–11	64% (7.45)
		1–22	64% (5.43)
Arbitrary	No Restrictions	1	47% (−1.85)
		1–11	52% (1.19)
		1–22	53% (1.21)

Note: This test is only valid to the extent that it represents a random sample.

[a] This is the percentage of squared forecast errors using restriction set A, which are smaller than the squared forecast errors using restriction Set B. There are a total of 1,908 unique forecast comparisons from the three models, six food groups, and three forecast horizons.

[b] This is a nonparametric sign test of the null hypothesis that the median-squared forecast errors are equal across both restriction types against the alternative hypothesis they are different from zero. A significantly positive statistic indicates that restriction set A has a lower median forecast error. Asymptotically and under the null hypothesis, the test statistic is distributed N (0, 1).

Tables 2 and 3 are useful for comparing mean-squared errors, but one may also wish to compare median-squared errors. The percentage of time that one restriction forecasts better than another is shown in Table 3 for various forecast horizons. Regardless of the forecast horizon, the economic restrictions provide smaller forecast errors than the arbitrary restrictions approximately 64% of the time. This percentage is significantly greater than 50% at all horizons. The difference in the percentage of times arbitrary restrictions outperform no restrictions is not significantly different from 50% but does seem to grow larger with the forecast horizon. This shows that, although mean-squared errors are significantly lower when arbitrary restrictions are used instead of no restrictions, there is little difference in median-squared errors.

Discussion

These results show unambiguous support for the use of economic restrictions in demand systems. Even though they are typically rejected in sample, the value of economic information is nicely demonstrated by their ability to improve forecasts. Part of this improvement is due to greater degrees of freedom, as seen by our results showing that restrictions can improve forecasts even when they are not true. However, much of this improvement emanates from the fact that relationships implied by economic theory are reflected in economic data, because demand systems with the symmetry condition consistently outperformed models with arbitrary restrictions or no restrictions.

Economists often make predictions using a combination of historical data and theory. According to the results of this study, the relative contribution of theory towards forecast accuracy grows as the forecast horizon lengthens. However, we should caution that these results pertain only to the symmetry restriction that is derived from the representative consumer model. Whether other restrictions, such as homogeneity, and other economic models can also improve forecasts is uncertain.

The symmetry restriction in its strict form is unlikely to be true, for the simple reason that no mathematical model is perfectly true and for this reason may be rejected in hypothesis tests. However, the symmetry restriction reflects basic relationships that seem to be reasonable even if consumers have different utility functions. Regardless of whether the specified demand function is correct or incorrect, the fact that beef is a strong substitute for pork suggests that pork should be a strong substitute for beef. As this study shows, this relationship becomes especially useful when predicting far into the future. The homogeneity condition reflects the simple fact that consumers face a budget constraint, so it likely contains nonsample information that is useful for forecasting as well.

Results here demonstrate that economic theory should not be tested solely by in-sample hypothesis tests. The ability of economic restrictions to significantly improve forecasts from demand systems demonstrates that the representative consumer theory is a valid theory, or at the least, a useful theory.

[Received November 2003; Accepted March 2004.]

References

Ashley, R., C.W.J Granger, and R. Schmalensee. "Advertising and Aggregate Consumption: An Analysis of Causality." *Econometrica* 48(1980): 1149–67.

Brieman, L. "The Little Bootstrap and Other Methods for Dimensionality Selection in Regression: X-Fixed Prediction Error." *Journal of the American Statistical Association* 87(1992):738–54.

Deaton, A., and J. Muellbauer. "An Almost Ideal Demand System." *American Economic Review* 70(1980):312–26.

Kastens, T.L., and G.W. Brester. "Model Selection and Forecasting Ability of Theory-Constrained Food Demand Systems." *American Journal of Agricultural Economics* 78,2(1996):301–12.

Mendenhall, W., D.D. Wackerly, and R.L. Schaeffer. *Mathematical Statistics With Applications,* 4th ed. Boston: PWS-Kent Publishing Company, 1990.

Sawa, T. "Information Criteria For Discriminating Among Alternative Regression Models." *Econometrica* 46,6(1978):1273–91.