Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 767

June 2003

IT Investment and Hicks' Composite-Good Theorem: The U.S. Experience

Jaime Marquez and Shing-Yi Wang

IT Investment and Hicks' Composite-Good Theorem: The U.S. Experience

Jaime Marquez and Shing-Yi Wang*

Abstract: We study whether aggregation residuals in U.S. private investment in information technology (IT) exhibit a predictable pattern that is consistent with Hicks' composite-good theorem and that may be used for forecasting. To determine whether one can extract such a pattern, we apply the general-to-specific strategy developed by Krolzig and Hendry (2001). This strategy combines ordinary least squares with a computer-automated algorithm that selects a specification based on coefficients' statistical significance, residual properties, and parameter constancy. Then, we derive the testable implications from Hicks' theorem and evaluate them with econometric formulations; we find qualified support for these implications. Having obtained these formulations, we evaluate their ex-post predictive accuracy and compare it to that of an autoregressive model. The key finding is that ignoring movement in relative prices results in a loss of information for predicting aggregation residuals.

Keywords: Aggregation Errors, Fisher Aggregates, Divisia Aggregate, General-to-Specific.

1   Introduction

The composite good theorem, postulated first by Hicks (1936), states that a collection of commodities can be treated as a single commodity when their prices move in parallel:[1]

> A collection of physical things can always be treated as if they were divisible into units of a single commodity so long as their relative prices can be assumed unchanged, in the particular problem in hand. J. R. Hicks, 1936, page 33.

In theory, then, changes in relative prices imply that one cannot treat units of real consumption and units of real investment as though they were divisible units of a single commodity called real GDP. Neither can one treat different types of investment–structures and equipment–as though they were divisible units of a single commodity called aggregate investment. The format of the U.S. national accounts recognizes explicitly the potential for this lack of additivity by using Fisher aggregation and reports the difference between the Fisher aggregate and the sum of its components; we term this difference the aggregation residual.[2] In practice, aggregation residuals for several decompositions of real GDP are small (figure 1); the largest of these residuals has a mean of a tenth of a percent of real GDP. Aggregation residuals for aggregate consumption, exports, and imports are also small. But for investment, the aggregation residual reaches eight percent of aggregate investment by 2003.

The patterns of these residuals raise interesting questions. For example, what is so unique about investment that it shows large aggregation residuals whereas the other macroeconomic aggregates do not? Further, how is it that changes in relative prices induce the "inverted-U" pattern observed in figure 1? To address these questions, we compute aggregation residuals for each of the components of investment and find that aggregation residuals in information-technology investment account for the patterns of figure 1. Having found the source of these residuals, we exploit Hicks' theorem to model them in terms of movements of relative prices and examine whether the parameter estimates explain the inverted-U pattern. Finally, we evaluate ex-post forecast performance in terms of confidence bands and compare them to those from an autoregressive formulation. We find that forecasts of the model based on relative prices are more accurate than forecasts of a model that ignores information about relative prices.

---

[1] For a formal treatment of this theorem, see Deaton and Muellbauer (1980, section 5.1).

[2] For Fisher aggregation in the U.S. national accounts, see Triplett (1992); Landefeld and Parker (1995); Varvares, Prakken, and Guirl (1998); Bachman, Jaquette, Karl, and Rocco (1998); and Whelan (2000).

2 Relative Prices and Aggregation Residuals

2.1 Data Considerations

To isolate the source of aggregation residuals of investment, we compute the aggregation residuals of its components: Non-residential Structures, Equipment and Software, and Residential; figure 2 shows these calculations. The data point to small aggregation residuals except for investment in Equipment and Software. Thus we disaggregate this investment category further into Industrial Equipment, Transportation Equipment, and Information Processing Equipment and Software (IT investment for short). The data indicate large aggregation residuals for IT investment (figure 2, bottom left panel) and its evolution closely resembles the inverted-U pattern of the residual for aggregate investment (compare the bottom right panels of figures 1 and 2). This consideration is important because spending on IT investment accounted for nearly a third of aggregate private investment spending, a share not matched by IT spending in other macroeconomic aggregates.

2.2 Analytical Considerations

For our analysis, we denote the Bureau of Economic Analysis (BEA) Fisher's aggregate as Y, the associated jth component as Y_j, and the number of components in the account as M. The composite-good theorem states that if relative prices are changing then $Y_t \neq \Sigma_j Y_{jt}$ and the aggregation residual R_t reported in the account reconciles the aggregate with the sum of its components: $Y_t = \Sigma_j Y_{jt} + R_t$.

To motivate how instability of relative prices gives rise to R, we use a simple example involving two aggregation methods. The first method is the Divisia approximation to the Fisher aggregate; the second method uses linear aggregation. The growth rate of the Divisia aggregate for Y is

$$\frac{dY_t}{Y_t} \equiv \widehat{Y}_t = \sum_j w_{jt} \widehat{Y}_{jt}, \tag{1}$$

where $w_{jt} = \frac{P_{jt} Y_{jt}}{\sum_s P_{st} Y_{st}}$ is the nominal expenditure share of the jth component and P_{jt} is the price of the jth component. The linear aggregate is $Y_t^\ell = \sum_j Y_{jt}$, and its growth rate is

$$\frac{dY_t^\ell}{Y_t^\ell} \equiv \widehat{Y}_t^\ell = \sum_j \frac{Y_{jt}}{Y_t^\ell} \widehat{Y}_{jt}, \tag{2}$$

2

where $\frac{Y_{jt}}{Y_t^\ell}$ is a real expenditure share. Equation (2) assumes that relative prices are fixed and thus, according to Hicks' theorem, embodies perfect substitutability among the $Y_j's$. Equation (1) does not assume that relative prices are fixed and thus substitutability among the $Y_j's$ could be less than perfect.

If we assume that prices move in parallel–that is, if $\frac{P_{jt}}{P_t} \equiv X$, then

$$\widehat{Y}_t = \sum_j w_{jt}\widehat{Y}_{jt} = \sum_j \left(\frac{XP_tY_{jt}}{\sum_s XP_tY_{st}}\right)\widehat{Y}_{jt} = \sum_j \left(\frac{Y_{jt}}{Y_t^\ell}\right)\widehat{Y}_{jt} = \widehat{Y}_t^\ell,$$

which means that the growth rates implied by the two methods are identical. Thus, applying the same initial condition to both methods, $Y_0^\ell = Y_0$, implies that there is no aggregation residual: $Y_t = \sum_j Y_{jt}$. To the extent that prices do not move in parallel, their movements will affect nominal shares and real shares differently. Specifically, a change in P_j affects nominal shares directly through valuation effects and indirectly through behavioral responses in the $Y_j's$. In contrast, a change in P_j has no valuation effect on real shares; these change only through the response of the $Y_j's$. As a result, changes in relative prices result in different growth rates for methods (1) and (2), a difference giving rise to R.

3   Empirical Framework

3.1   Model Specification

We postulate that aggregation residuals behave according to

$$\frac{R_t}{Y_t} = \rho_t + v_t, \tag{3}$$

where ρ_t is the aggregation residual per dollar, $v_t = \frac{u_t}{1-\alpha(L)}$, $u_t \tilde{} N(0, \sigma^2)$, and $\alpha(L)$ is a polynomial in the lag operator L. This formulation allows for the serially correlated pattern in the data and it recognizes that the importance of R_t's magnitude hinges not so much on its absolute value but on its value relative to the account's own aggregate. We scale R using the Fisher aggregate to avoid internal consistencies that would arise from using an alternative aggregate that assumes that relative prices are fixed when in fact they are not.

Following Hicks' theorem, we model ρ_t in terms of relative prices. Many such prices

are suitable candidates but, given the smallness of our sample, we consider just four: the real interest rate, the relative price of oil, the real exchange rate, and the relative price of computers. The real interest rate captures the role of intertemporal substitution; the real price of oil captures substitutability between energy and non-energy productive factors; the real exchange rate captures substitutability between tradeable and non-tradeable products; the relative price of computers captures the substitutability between IT-capital and other forms of capital. Sole reliance on the role of substitution effects assumes, however, that substitution possibilities are fixed over time: there is no innovation. If there are innovations, then measurement should be based on reservation, instead of market, prices (Fisher and Shell, 1972, p. 104). Because reservation prices are not observed, we model the effects of innovation with a trend and postulate that

$$\rho_t = \rho(X_{1t}, ..., X_{4t}, \ Trend).$$

In the absence of an accepted theory for specifying $\rho(\cdot)$, we postulate an autoregressive distributed lag allowing for nonlinearities:

$$[1 - \alpha(L)]\frac{R_t}{Y_t} = \beta + \gamma Trend + \sum_i \phi_i(L)X_{it} + \sum_i \varphi_i(L)X_{it}^2 + u_t, \tag{4}$$

where $\phi_i(L)$ and $\varphi_i(L)$ are polynomials in the lag operator L. For equation (4) to be consistent with Hicks' theorem, relative prices should have only transitory effects on R. Finding that relative prices have permanent effects would contradict this theorem.

3.2   Forecast Applicability

If one retains the format of the national accounts, then our approach helps forecast the Fisher aggregate Y in terms of forecasts of its parts:

$$Y_t^f = \sum_{j=1}^{M} Y_{jt}^f + R_t^f, \tag{5}$$

where " f " denotes a forecasted value. An alternative approach involves using a Divisia approximation to estimate BEA's Fisher aggregate; the resulting estimate is $Y_t' = Y_{t-1}'(1 + \sum_{j=1}^{M} w_{jt}\widehat{Y}_{jt})$. Reliance on such estimates will not replicate, even in principle, BEA's published

4

value of Y because they use as inputs the reported values of $Y_1, ..., Y_M$ instead of the raw data used by the BEA to measure Y. As a result, reconciling BEA's Y with the estimate Y' involves an aggregation residual $R' : Y_t = Y'_t + R'$ and this residual needs to be forecasted along with the other variables used in the Divisia approximation:

$$
\begin{aligned}
Y_t^f &= Y_t^{f\prime} + R_t^{f\prime} \\
Y_t^{f\prime} &= Y_{t-1}^{f\prime}(1 + \sum_{j=1}^{M} w_{jt}^f \, \widehat{Y}_{jt}^f) \\
w_{jt}^f &= \frac{P_{jt}^f Y_{jt}^f}{\sum_s P_{st}^f Y_{st}^f}.
\end{aligned}
\tag{6}
$$

Equation (6) is appealing because it mimics the aggregation procedures used by the BEA. However, implementing equation (6) needs forecasts of expenditure shares which in turn involves forecasting investment prices. Thus whether the forecast errors from equation (6) are lower than those of equation (5) is an unresolved question deserving further study.

4   Empirical Results

4.1   In-Sample Estimates: Parameters and Tests

We use equation (4) to explain movements of aggregation residuals of IT investment; the data are quarterly and span from 1987:Q1 to 2002:Q3. For the relative price of oil (p) we use the logarithm of the ratio of the unit-value of U.S. oil imports to the price index of investment (figure 3); for the relative price of computers and software (w), we use the logarithm of the price index for investment in computers, peripherals, and software relative to the price index of investment. For the real interest rate (r) we consider three measures that differ in their maturity: the real federal funds rate; the real 10-year Treasury rate, and the real 30-year Treasury rate; all of these ex-post rates are constructed as the associated nominal rate minus the CPI inflation lagged one quarter. For the real effective exchange rate (q) we use two measures: the logarithm of the Federal Reserve's *Broad* real exchange rate, based on industrial and developing countries' currencies; and the logarithm of the IMF's real effective exchange rate, based on industrial countries' currencies only. Using these variables, we assemble six models that differ in their measures of real interest and exchange rates.

For estimation, we assume a maximum number of five lags for the dependent variable and one lag for each of the remaining explanatory variables; we also include a "9/11" dummy variable equal to one in the third quarter of 2001 and zero otherwise. The usefulness of our tests increases with the elimination of irrelevant variables and thus we apply the general-to-specific strategy developed by Krolzig and Hendry (2001). This strategy combines ordinary least squares with a computer-automated algorithm that selects a specification in four stages:

1. Estimating the parameters of an unrestricted formulation–equation (4)–and testing for congruency (white-noise residuals and parameter constancy).[3]

2. Implementing multiple reduction paths simultaneously. One reduction path could get started by excluding the least significant variable whereas another reduction path could get initiated by excluding a block of variables that are statistically insignificant.

3. Testing whether the specification from a reduction path is congruent. If it is, then implement another round of reductions and test for congruency; continue this process until the specification violates congruency. In that case, the algorithm selects the immediately prior specification and labels it *Final model*.

4. Collecting the Final models from various reduction paths and applying encompassing tests to them. The specification that encompasses all others becomes the *Specific model*. If there is no single encompassing model, then the algorithm forms a "union" model using the variables from all of the Final models and re-starts the specification search from step (2). If this strategy fails to yield a single Specific model, then the algorithm applies three information criteria (Akaike, Schwarz, and Hannan-Quinn) to the Final models and selects the one that minimizes all these criteria; that model becomes the Specific model. Otherwise, the algorithm fails to find a Specific model.

A key feature of this algorithm is the sequential adjustment of significance levels to recognize the joint nature of model specification and parameter estimation.

Table 1 reports the test results and indicates which variables have statistically significant effects. In terms of coefficient estimates, relative prices have short-run, nonlinear effects on

---

[3]To test normality we use the Jarque-Bera statistic which is distributed as $\chi^2(2)$; to test serial independence we use an F-test for the hypothesis that the coefficients for an AR(4) of the estimation residuals are jointly equal to zero; to test homoskedasticity we use an ARCH test. For parameter constancy, we implement two Chow tests: one excluding half of the sample and another excluding the last eight observations.

the aggregation residual of IT investment, a finding that supports Hicks' composite-good theorem.[4] Relative prices, however, have permanent effects which is not consistent with Hicks' theorem. Two factors may account for these permanent effects. First, our selection of relative prices could be incomplete and thus yields biased coefficients. Second, changes in relative prices, especially those associated with IT products are inducing permanent changes on substitution possibilities whereas Hicks' theorem assumes that substitution possibilities are fixed. Finally, the trend has a significant effect on the aggregation residual in four of the six models; this finding suggests that the innovation process is affecting substitution possibilities that are not captured by the choice of relative prices in those four models.

To assess the models' ex-post predictive accuracy, we first re-estimate their coefficients excluding the last eight observations–2000:Q4 to 2002:Q3–and then implement one-step ahead and multi-step ahead simulations; these simulations use the observed data for relative prices. The root mean squared error (RMSE) for one-step ahead simulations ranges from 1.0% (M1) to 1.4% (M5); for multi-step ahead simulations, the RMSE ranges from 1.6% for model M3 to 3.7% for model M5.[5]

Overall, model M3 has the highest ex-post predictive accuracy and explains the data well (figure 4, top panel). Moreover, in terms of its statistical properties, the trend has no significant effect, one cannot reject the assumptions made for the disturbances and, importantly, one cannot reject the hypothesis of parameter constancy. Evidence of parameter constancy allows us to use the coefficient estimates to assess the source of the inverted-U pattern of IT investment aggregation residuals. To this end, figure 4 (bottom panel) shows the effect of the relative price of IT products (w) on the aggregation residual (R/Y) using the observed data for w. The calculation shows a positive effect of w on (R/Y) that grows over time and reaches a maximum around 1997. Overall, the profile of this effect closely resembles that of the aggregation residual of IT investment in figure 2. Determining whether the non-linear effects of w are the sole reason for this inverted-U is beyond the scope of this paper but the results suggest that nonlinearities are part of the answer.

---

[4]The coefficient estimates are in appendix 1, available in http://www.federalreserve.gov/pubs/ifdp/2003/767/

[5]The forecasts are in appendix 2, available in http://www.federalreserve.gov/pubs/ifdp/2003/767/

4.2 Out-of-Sample Predictions: Confidence Bands

Figure 5 shows the 95 percent confidence bands for the ex-post predictions from model M3 over 2000:Q4-2002:Q3. For 1-step ahead simulations, the model has statistically significant underpredictions in 2001:Q2 and 2001:Q3 (figure 5, top left panel); these underpredictions increase when one uses multi-step simulations (figure 5, top right panel). Thus, the effects from the "9/11" episode are not easily predicted by model M3. Yet, the model's predictions return to the actual path and, indeed, the last three prediction errors are not significant.

To assess whether movements in relative prices are informative for forecasting, figure 5 also reports the 95 percent confidence bands for the predictions from an autoregressive model that uses five lags and a constant. Comparing the one-step ahead predictions across the two models suggests that ignoring movements in relative prices results in no loss of information: the RMSE are virtually identical for the two models. But for multi-step ahead simulations, the time-series model shows a growing prediction error which translates into a higher RMSE: three percent relative to 1.6 percent for the relative-price based model. These results suggest that ignoring movements in relative prices results in a loss of information for predicting aggregation residuals in IT investment.

5 Conclusions

This paper studies whether aggregation residuals in IT investment exhibit a pattern reliable enough to be statistically characterized and used for forecasting. To determine whether one can extract a pattern from those residuals, we implement several regression models and test whether the parameter estimates are consistent with Hicks' composite-good theorem. We find that changes in relative prices help explain short-run movements in aggregation residuals. We also find that changes in relative prices have permanent effects on aggregation residuals which contradicts a key prediction of Hicks' composite-good theorem. One factor that may account for these permanent effects is that our selection of relative prices is incomplete. Another factor is that the product mix of information-technology products is changing over time due to innovation whereas Hicks' theorem assumes an unchanged product mix.

With this caveat in mind, we implement ex-post simulations and find statistically significant underpredictions for aggregation residuals in IT investment during the "9/11" episode. These underpredictions are temporary, however, and the predictions return to the actual

path. Moreover, using an autoregressive model as an alternative yields larger prediction errors suggesting that ignoring movement in relative prices results in a loss of information.

6   References

Bachman, D., P. Jaquette, K. Karl, and P. Rocco, 1998, "The WEFA U.S. Macro Model with Chain-Weighted GDP," *Journal of Economic and Social Measurement*, 24, 143-155.

Deaton, A. and J. Muellbauer, 1980, *Economics and Consumer Behavior*, Cambridge: Cambridge University Press.

Fisher, F. and K. Shell, 1972, *The Economic Theory of Price Indices*, Academic Press, New York.

Hendry, D. F. and H. Krolzig, 2001, *Automatic Econometric Model Selection Using PcGets*, London: Timberlake.

Hicks, J.R., 1939, *Value and Capital*, Oxford: Oxford University Press (1978 printing, second edition).

Krolzig, H. and D.F. Hendry, 2001, "Computer Automation of General-to-Specific Model Selection Procedures," *Journal of Economic Dynamics & Control*, 25, 831-866.

Landefeld, S. and R. Parker, 1995, "Preview of the Comprehensive Revision of the National Income and Product Accounts BEA's New Featured Measures of Output and Prices," *Survey of Current Business*, July, 31-38.

Triplett, J., 1992, "Economic Theory and BEA's Alternative Quantity and Price Indexes," *Survey of Current Business*, April, 49-52.

Varvares, C., J. Prakken, and L. Guirl, 1998, "Macro Modeling with Chain-type GDP," *Journal of Economic and Social Measurement*, 24, 123-142.

Whelan, K. 2000, "A Guide to the Use of Chain Aggregated NIPA Data," Finance and Economics Discussion series 2000-35; Washington, D.C. : Board of Governors of the Federal Reserve System.
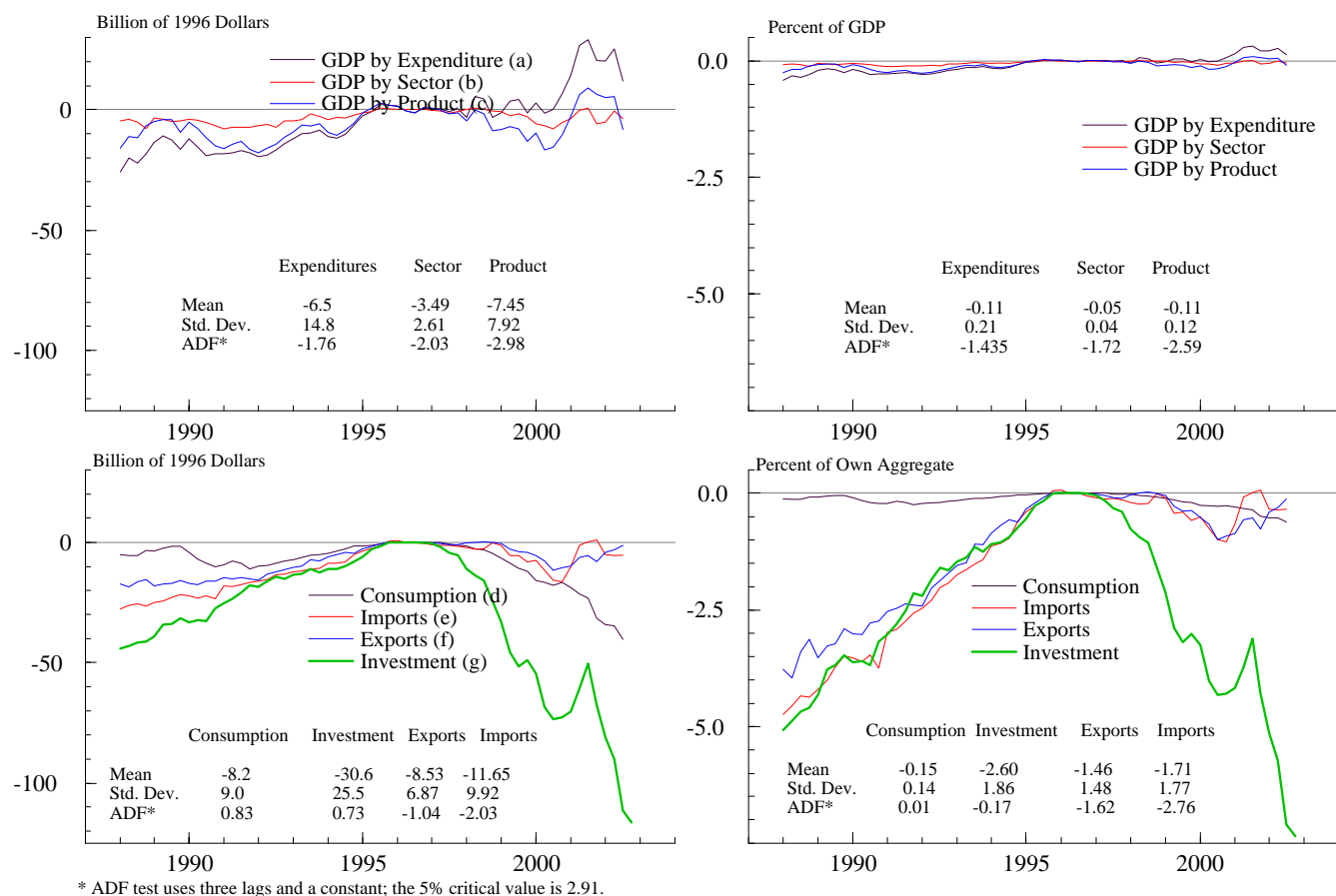
Figure 1: Aggregation Residuals in U.S. National Income and Product Accounts. *Source: Survey of Current Business.*

[a] Aggregation residual for GDP by type of Expenditure (NIPA table 1.2).

[b] Aggregation residual for GDP by type of Sector (NIPA table 1.7).

[c] Aggregation residual for GDP by type of Product (NIPA table 1.4).

[d] Aggregation residual for Consumption (NIPA table 2.3).

[e] Aggregation residual for Imports of Goods and Services (NIPA table 4.4).

[f] Aggregation residual for Exports of Goods and Services (NIPA table 4.4).

[g] Aggregation residual for Private Fixed Investment (NIPA table 5.5).
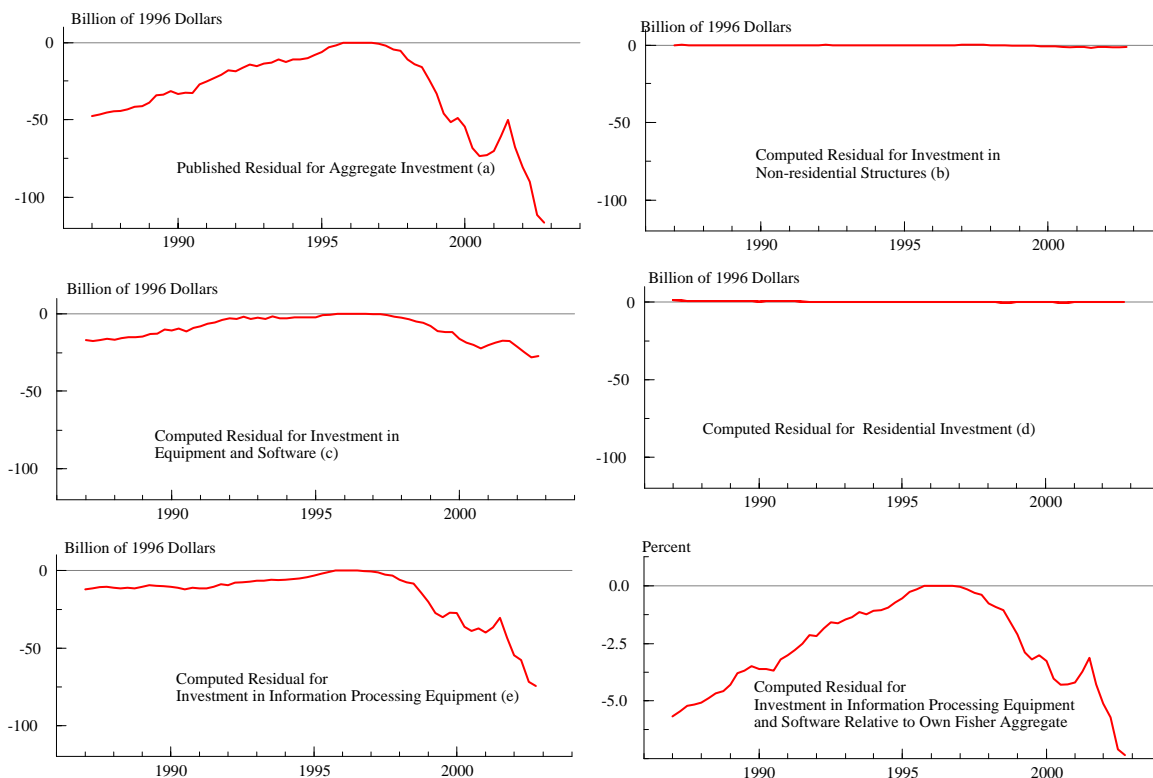
Figure 2: Aggregation Residuals in Aggregate Private Investment. *Source: Survey of Current Business.*

[a] Aggregation residual for Private Investment (last line of NIPA table 5.5).

[b] Computed by authors as the difference between Private Investment in Structures and the sum of expenditures on Non-residential Buildings, Utilities, Mining Exploration, Shafts and Wells, Other Non-residential Structures.

[c] Computed by authors as the difference between the Fisher aggregate for Private Investment of Equipment and Software and the sum of Information Processing Equipment and Software, Industrial Equipment, Transportation Equipment, Other (Equipment and Software).

[d] Computed by authors as the difference between the Fisher aggregate for Private Investment in Residential and the sum of Single-family Structures, Multifamily Structures, Other Residential Structures, and Residential Equipment.

[e] Computed by authors as the difference between Fisher aggregate for Private Fixed Investment in Information Processing Equipment and Software and the sum of Investment in Computers and Peripheral Equipment, Software, and Other Information Processing Equipment and Software.
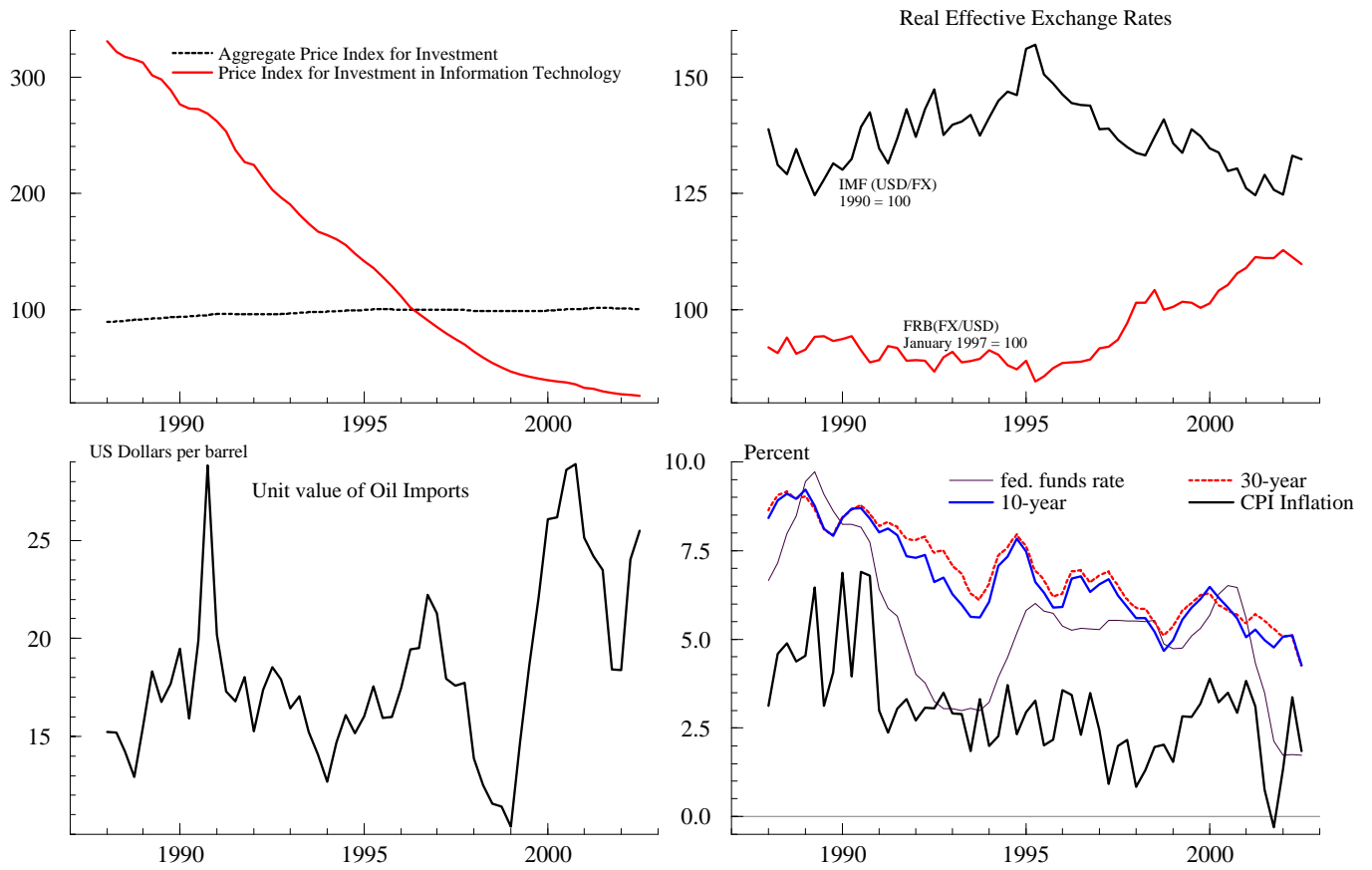
11

Figure 3: Evolution of Explanatory Variables. *Source: Survey of Current Business, Bureau of Labor Statistics, Intenrational Monetary Fund, and Federal Reserve Board.*
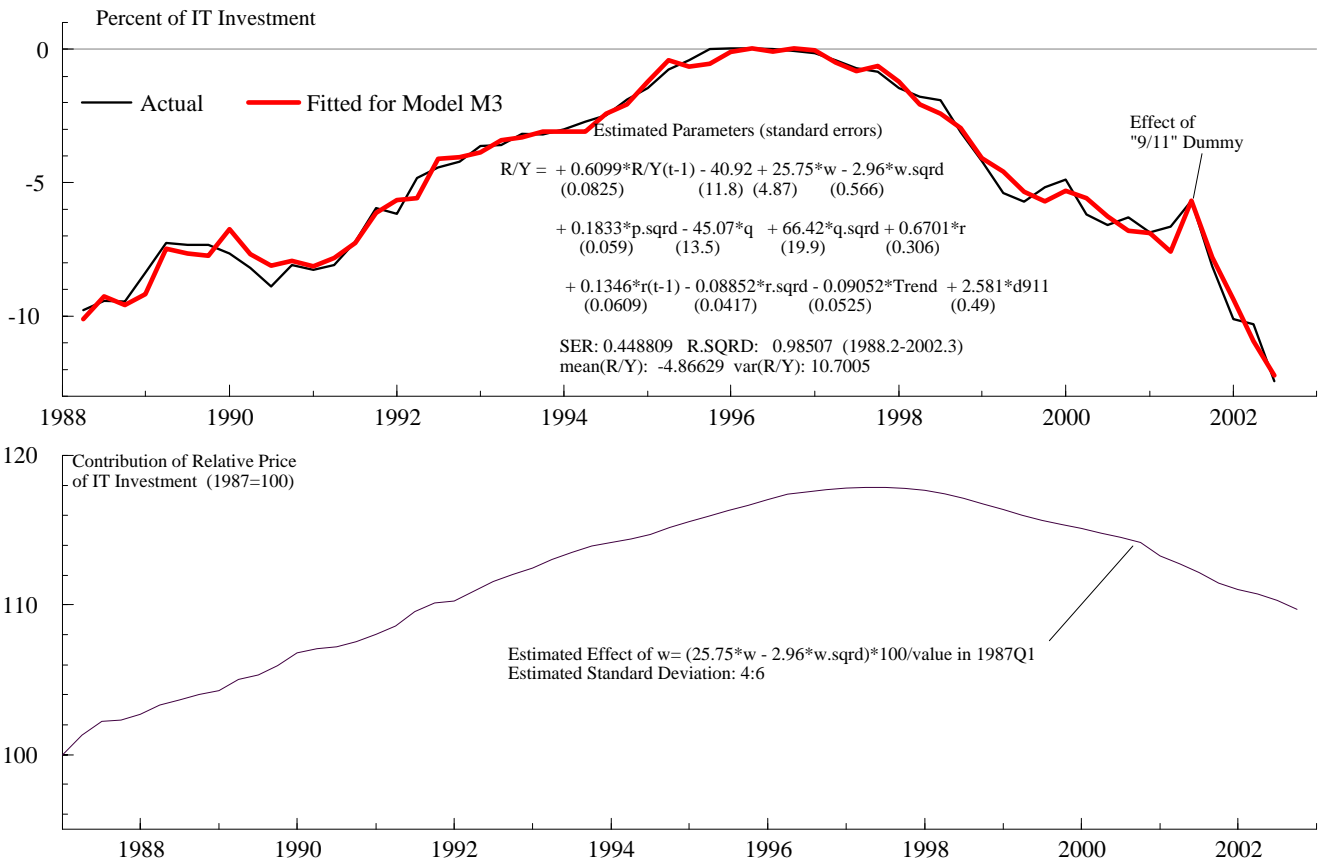
Figure 4: Empirical Results for Model M3-Fitted Values and Effect of Relative Price of IT Investment
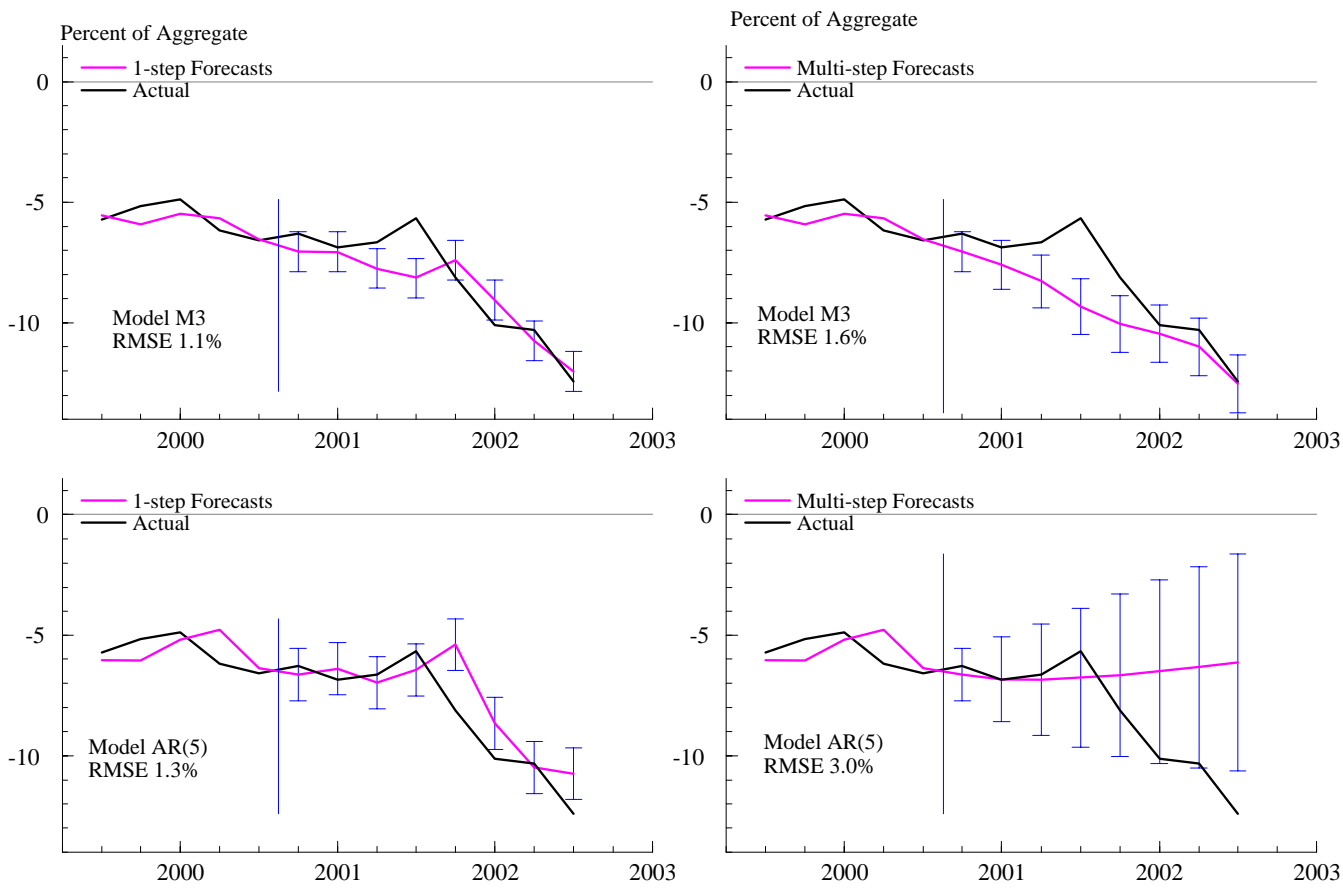
Figure 5: 95% Confidence Intervals for Ex-post Forecasts of Residuals in IT Investment

Table 1: Parsimonious Model of Aggregation Residuals Private Investment in Information Equipment [a]
Ordinary Least Square Estimation Results - 1987:Q1-2002:Q3
Sensitivity to Model Specification [b]

	M1	M2	M3	M4	M5	M6
Significant short-run price effects						
Linear	w, p, q, r	w, p, q	w, q, r	w, q	w, q	w, q
Nonlinear	w	w, p, q	w, p, q, r	w, p, q	w, p, q	w, p, q
Significant long-run price effects						
Linear	w, r	w, p, q	w, q, r	w, q	w	w, q
Nonlinear	w	w, p, q	w, p, q	w, q	w, q	w, q
Is Trend Significant?	no	yes	no	yes	yes	yes
Adj. R^2	0.98	0.98	0.99	0.98	0.98	0.98
SER (percent)	0.46	0.45	0.45	0.45	0.47	0.45
Number of Parameters	11	14	12	13	11	13
Lagged Dep. Variable: $\alpha(1)$	0.56	0.50	0.61	0.58	0.62	0.58
(std. error)	(0.09)	(0.10)	(0.08)	(0.09)	(0.09)	(0.09)
Hypothesis Testing [c]						
Properties of Disturbances						
Normality	0.62	0.51	0.91	0.43	0.51	0.44
Serial Correlation	0.82	0.31	0.50	0.87	0.60	0.87
Homoskedasticity	0.14	0.84	0.56	0.40	0.48	0.40
Parameter Stability						
Exclude Half Sample	0.09	0.32	0.62	0.42	0.10	0.42
Exclude Last Eight Obs.	0.10	0.09	0.06	0.02*	0.14	0.02*
Forecasts Error (RMSE) [d]						
Multi-step Ahead	1.78	2.01	1.64	1.80	3.66	1.80
One-step Ahead	1.04	1.13	1.11	1.08	1.40	1.08

[a] Aggregation residual for Private Investment in Information Equipment (NIPA table 5.5): difference between the BEA aggregate for nonresidential, private investment in information equipment and the sum of investment in computers and peripheral equipment, software, and other information processing equipment and software.

q: Real effective exchange rate,
r: Real interest rate,
p: Price of oil relative to aggregate's own price
w: Price of computer equipment purchases relative to aggregate's own price

[b] Models differ in their choice of exchange-rate and interest-rate measures:

	Real Interest Rate		
Real Effective Exchange Rate	Federal Funds	10-year	30-year
IMF-Narrow	M1	M2	M3
FRB-Broad	M4	M5	M6

[c] Significance level needed to reject the associated hypothesis; an entry less than 0.05 means that one can reject the null hypothesis at the five percent significance level-denoted with an asterisk.

[d] Root mean squared error over 2000:Q4 to 2002:Q3.