

Do Data Revisions Matter for DSGE Estimation?

Gregory E. Givens^{a,*}

^a*Department of Economics, Finance, and Legal Studies, University of Alabama, Tuscaloosa, AL 35487, USA*

First Draft: April 2016

This Draft: June 2016

Abstract

This paper checks whether the coefficient estimates of a famous DSGE model are robust to macroeconomic data revisions. The effects of revisions are captured by rerunning the estimation on a real-time data set compiled using the latest time series available each quarter from 1997 through 2015. Results show that point estimates of the structural parameters are generally robust to changes in the data that have occurred over the past twenty years. By comparison, estimates of the standard errors are relatively more sensitive to revisions. The latter implies that judgements about the statistical significance of certain parameters depend on which data vintage is used for estimation.

Keywords: Data Revisions, Real-Time Data, DSGE Estimation

JEL Classification: C32, C82, E32, E52

*Corresponding author. Tel.: + 205 348 8961.

E-mail address: gegivens@cba.ua.edu (G.E. Givens).

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models have become popular tools for macroeconomic analysis. Part of the appeal is that they provide a complete multivariate representation for the data using a theoretical framework built strictly on optimizing principles and market-clearing assumptions. This aspect makes it possible for DSGE practitioners to interpret the various co-movements and response patterns seen in the data in terms of explicit microeconomic concepts. For example, inflation inertia, when viewed through a DSGE lens, might be interpreted as evidence of staggered price and wage setting. The persistence in certain aggregate quantities, on the other hand, could be due to a mix of real frictions affecting the way households allocate consumption and capital over the business cycle.

Ordinarily DSGE models are estimated using the latest vintage of historical data that is available at the time. The data sets that were available to analysts in the past, however, often bear little resemblance to the newly published data sets on hand today. The reason is that time-series data for many aggregate price and quantity variables are continually revised in the months and years after their initial release as statistical agencies acquire better source information and as underlying measurement concepts evolve (e.g., Croushore and Stark, 2001; Croushore, 2011). A natural question then is whether estimation results made famous in the past hold up to changes in the historical series caused by subsequent data revisions. Stated differently, are the leading interpretations of key macroeconomic relationships, as seen through the prism of a contemporary DSGE model, robust to data revisions?

I try to answer this question by re-estimating a well-known DSGE model using historical data as it existed at different points in time. The main objects of interest here are estimates of the behavioral parameters describing the preferences and technologies of households and firms. The influence that data revisions have on these objects can be found by scrutinizing how the estimates vary as data are drawn from consecutive vintages while holding the sample period fixed throughout.¹ The basic findings of the study are considered robust if the parameter estimates are relatively stable across vintages. If instead the estimates turn out to be fragile, then the model's structural interpretation of the data may be of limited value.

The empirical exercise sketched above borrows heavily from Croushore and Stark (2003) and Croushore and Evans (2006). Both sets of authors rerun several prominent macroeconomic studies using different vintages of the data. What separates my work from theirs is the class of model chosen for testing. All of the replication experiments conducted by

¹By "vintage" I mean the version of the data that was reported on a specific date (e.g., Swanson, 1996).

Croushore, Stark, and Evans involve “reduced-form” models depicting relationships among variables based loosely on economic theory.² In this paper I evaluate the robustness of a fully coherent business cycle model with intertemporal general-equilibrium foundations.

The particular DSGE model considered here is the one developed by Christiano, Eichenbaum, and Evans (2005). Their model adds a number of real and nominal frictions onto an otherwise standard neoclassical growth structure that enables it to capture regular features of the data more accurately. It is an ideal candidate to put to a vintage robustness test for many reasons. One is its age. The version that appeared in print has actually been around for a long time and was originally estimated using sample data that runs through 1995. The multitude of revisions that have occurred since then should provide the researcher with ample information about their effects on parameter inference. Another reason is that most state-of-the-art DSGE models inherit core components from the Christiano-Eichenbaum-Evans framework (e.g., Smets and Wouters, 2007; Altig, Christiano, Eichenbaum, and Lindé, 2011). Thus any conclusions drawn from this study about the significance of data revisions might also apply to the latest crop of DSGE models.

A comprehensive analysis of the Christiano-Eichenbaum-Evans model indicates that point estimates of the behavioral parameters are generally robust to data revisions. Using historical data as it appeared in every quarter from 1997 through 2015, I find that the amount of variation in the estimates across vintages, as measured by standard deviation, is noticeably smaller than the typical *sampling* variability of the estimates from any given vintage. This suggests that the common practice of looking only at the most recent data sets may be sufficient for researchers seeking valid estimates of empirical DSGE models.

When it comes to the assessment of statistical significance, however, this strategy may not be as robust for the simple reason that standard error estimates are, by comparison, less stable than the corresponding parameter estimates. As it turns out, three of the seven estimated parameters experience abrupt shifts in their standard errors around dates where major historical revisions are known to have occurred. In each case the outcome of a formal hypothesis test that the true parameter value equals zero depends on whether the vintage precedes or follows the revision date. These results suggest that DSGE practitioners should probably exercise caution when drawing conclusions about parameter significance based solely on a single vintage of macroeconomic data.

In the course of testing for robustness to data vintage, it is critical that the researcher

²Included are the identified VAR studies of Blanchard and Quah (1989), Christiano, Eichenbaum, and Evans (1996), and Galí (1992), the regression analysis of the life-cycle-permanent-income hypothesis from Hall (1978), and the cross-correlation statistics reported in Kydland and Prescott (1990).

distinguish between two types of revisions. As explained by Croushore and Stark (2003), revisions can be *information-based* or *structural*. The former occurs when government statistical agencies update earlier estimates with the regular arrival of more complete source information, a process that normally unfolds in the first few years after an initial data release. The latter takes place only when they implement wholesale changes to the way the data are constructed. Examples include things like advances in aggregation methods, base year adjustments, or definitional changes to the measurement concepts themselves.

To be clear, this paper focuses primarily on testing whether structural data revisions matter for DSGE estimation. Less scrutiny is given here to information-based revisions. That being said, it is impossible to perfectly isolate the two because some vintages actually contain elements of both.³ But by holding the sample period constant across vintages, I ensure that the most prominent revisions implicit in the real-time data set are mainly of the structural sort. This follows from the fact that structural revisions largely dominate all but the earliest vintages considered. Information-based revisions, because they mostly affect just the latest sample observations, lose influence as time goes on.

1.1 Related Studies

There is to date a mature literature that deals with the implications of data revisions for macroeconomic forecasting and policy analysis (e.g., Croushore, 2006; Orphanides, 2001; Nikolsko-Rzhevskyy, 2011). By contrast, relatively little work has been done questioning how revisions might affect the estimation of structural models, particularly those with general equilibrium characteristics.⁴ Notable exceptions are the studies of Tetlow and Ironside (2007) and Tetlow (2012). Central to both papers is an analysis of successive vintages of FRB/US, the Federal Reserve Board's large-scale model of the US economy. A common finding is that estimates of key policy multipliers implied by the model have fluctuated considerably over time (i.e., across vintages). Although related, this research program differs from my work in some important ways. First, FRB/US is not a pure DSGE model because only a subset of the behavioral equations reflect the decision rules of optimizing agents. Second, the changes in model properties observed by the authors are not necessarily attributable to data revisions alone. The reason is that the estimation period for FRB/US has varied from vintage to vintage, and whole portions of the model describing activity in various sectors have been

³An example of such an occurrence is when one series in a data set undergoes an information-based revision while a separate series undergoes a structural revision in the same quarter.

⁴In a recent study Croushore and Sill (2014) use an empirical DSGE model to show that the fundamental shocks driving ex-post revised data also influence, to a large extent, the dynamics of real-time data.

added or dropped along the way. Both of these factors in conjunction with data revisions are responsible for the apparent lack of robustness in estimates of the policy multipliers.

2 The Christiano-Eichenbaum-Evans Model

The equilibrium model described in Christiano *et al.* (2005) contains several real and nominal frictions that affect the spending and allocation decisions of households and firms. The nominal frictions include staggered price and wage contracts à la Calvo (1983) and Yun (1996), augmented with full indexation to past inflation, along with a transactions role for cash balances via money-in-the-utility-function and a working capital constraint on production. The real rigidities derive from imperfect competition in product and labor markets, internal habit formation in consumption, and costs of adjusting both investment and the utilization rate of capital. The model also has a financial intermediary that accepts cash deposits from households, makes loans to producers, and receives lump-sum injections from the monetary authority. The only source of uncertainty in the economy is a shock to monetary policy.

In the rest of this section, I comment briefly on the log-linearized version of the Christiano-Eichenbaum-Evans model that will later be estimated using alternative vintages of US data. The list of equations and variable definitions are presented in Table 1, where in the interest of uniformity, I have adopted the same notation used by the original authors. Because the model is already so familiar, my goal here is simply to call attention to the key parameters that appear in the various behavioral equations and which constitute the focus of this study. Table 2 contains a list of those parameters along with the values reported in Christiano *et al.* (2005). Readers should consult Christiano, Eichenbaum, and Evans (2001) and the appendix therein for an exposition of the full nonlinear general equilibrium model.

2.1 Behavioral Equations

Eqs. (M-1) and (M-2) follow from the household's first-order conditions for consumption and investment. Lagged consumption appears in (M-1) as a result of internal habit formation in the utility function. Parameter $b \in [0, 1)$ indexes the strength of the habit factor relative to current consumption. Households also face costs of transforming current and past investment into productive capital. This assumption is evident in (M-2), where $1/\kappa > 0$ determines the elasticity of investment with respect to a 1 percent rise in the price of installed capital.⁵

⁵Forward substitution on (M-2) yields $i_t = i_{t-1} + (1/\kappa) \sum_{j=0}^{\infty} \beta^j E_{t-1} P_{k',t+j}$.

Table 1**The log-linearized Christiano-Eichenbaum-Evans model**

Consumption demand	$(1 + \beta b^2)c_t = bc_{t-1} + \beta b E_{t-1} c_{t+1} - (1 - b)(1 - \beta b) E_{t-1} \psi_t$	(M-1)
Investment demand	$P_{k',t} = \kappa(i_t - i_{t-1}) - \beta \kappa(E_{t-1} i_{t+1} - i_t)$	(M-2)
Arbitrage equation	$P_{k',t} = \beta(1 - \delta) E_{t-1} P_{k',t+1} + (1 - \beta(1 - \delta)) E_{t-1} r_{t+1}^k - E_{t-1}(\psi_t - \psi_{t+1})$	(M-3)
Capital utilization	$\sigma_a u_t = E_{t-1} r_t^k$	(M-4)
Capital accumulation	$\bar{k}_{t+1} = (1 - \delta)\bar{k}_t + \delta i_t$	(M-5)
Money stock	$\psi_t = E_t \psi_{t+1} + E_t(R_{t+1} - \pi_{t+1})$	(M-6)
Cash balances	$q_t = -\frac{1}{\sigma_q} \left(\frac{\mu}{\mu - \beta} R_t + \psi_t \right)$	(M-7)
Cost minimization	$r_t^k - (w_t + R_t) = H_t - (u_t + \bar{k}_t)$	(M-8)
Real marginal cost	$s_t = \alpha r_t^k + (1 - \alpha)(R_t + w_t)$	(M-9)
Calvo price-setting	$\pi_t = \frac{1}{1 + \beta} \pi_{t-1} + \frac{\beta}{1 + \beta} E_{t-1} \pi_{t+1} + \frac{(1 - \beta \xi_p)(1 - \xi_p)}{(1 + \beta) \xi_p} E_{t-1} s_t$	(M-10)
Calvo wage-setting	$w_t + \pi_t = \frac{1}{1 + \beta} (w_{t-1} + \pi_{t-1}) + \frac{\beta}{1 + \beta} E_{t-1} (w_{t+1} + \pi_{t+1}) \dots$ $+ \frac{(1 - \beta \xi_w)(1 - \xi_w)}{(1 + \beta) \xi_w} \left(\frac{\lambda w - 1}{2\lambda w - 1} \right) (H_t - E_{t-1} \psi_t - w_t)$	(M-11)
Resource constraint	$\left(\frac{1/\beta - (1 - \delta) - \alpha \delta}{1/\beta - (1 - \delta)} \right) c_t + \left(\frac{\alpha \delta}{1/\beta - (1 - \delta)} \right) i_t + \alpha u_t = \lambda_f \left(\alpha(u_t + \bar{k}_t) + (1 - \alpha) H_t \right)$	(M-12)
Loan market clearing	$w_t + H_t = \frac{\mu}{\mu - q/m} (\mu_t + \mu_{t-1} + m_{t-1} - \pi_t) - \frac{q/m}{\mu - q/m} q_t$	(M-13)
Money growth	$m_t = \mu_{t-1} + m_{t-1} - \pi_t$	(M-14)
Monetary policy	$\mu_t = \theta_0 \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \theta_3 \epsilon_{t-3} + \dots$	(M-15)

Notes: All variables are expressed as log deviations from the deterministic steady state: c - consumption; ψ - shadow value of currency; $P_{k'}$ - shadow value of capital; i - investment; r^k - real rental rate; u - capital utilization rate; \bar{k} - capital stock; R - nominal interest rate; π - inflation rate; q - real cash balances; w - real wage; H - aggregate labor supply; s - real marginal cost; μ - growth rate of money; m - real money balances; ϵ - policy innovation.

Eq. (M-3) derives from an arbitrage condition linking the marginal cost of increasing the capital stock to the return on investment. The first-order condition for capital utilization leads to (M-4). It states that the marginal benefit of increasing the utilization rate, the real rental price of capital services, equals the additional resource costs concomitant with such an adjustment. Here $1/\sigma_a > 0$ is the elasticity of utilization with respect to a unit increase in the rental rate. Finally, the law of motion for capital is given by (M-5).

Eqs. (M-6) and (M-7) originate from the first-order conditions for next period's money stock and current nominal cash balances, which is the portion of the money supply not deposited with the financial intermediary. The latter condition ensures that households are indifferent between allocating a dollar to cash balances or the intermediary. The preference parameter $\sigma_q > 0$ determines the interest semielasticity of money demand, or the percent reduction in the demand for cash balances after a 1 percent rise in the nominal interest rate.

Every period intermediate good firms minimize their variable production costs. As indicated by (M-8), these efforts imply that the marginal rate of transformation between labor and capital services equals the relative factor price. Because each firm must borrow its wage

Table 2
Structural parameters

Parameter	Description	Value
β	household's subjective discount factor	$1.03^{-0.25}$
b	degree of internal habit formation	0.65^\dagger
κ	inverse elasticity of investment	2.48^\dagger
δ	depreciation rate of physical capital	0.025
σ_a	inverse elasticity of capital utilization	0.01^\dagger
σ_q	inverse interest semielasticity of money demand	10.62^\dagger
μ	gross steady-state growth rate of money	1.017
α	steady-state share of capital income	0.36
ξ_p	probability of not reoptimizing price contracts	0.60^\dagger
ξ_w	probability of not reoptimizing wage contracts	0.64^\dagger
λ_w	gross steady-state wage markup	1.05
λ_f	gross steady-state price markup	1.20^\dagger
q/m	steady-state fraction of money held as cash balances	0.44

Notes: † denotes a value that is estimated in Christiano *et al.* (2005). All others are fixed by the authors prior to estimation.

bill in advance of production, the cost of employing a unit of labor is the product of the real wage and the nominal interest rate. With capital-labor ratios identical across firms, real marginal cost depends only on the factor prices as illustrated in (M-9).

Eqs. (M-10) and (M-11) characterize how aggregate prices and wages are determined. In short, a continuum of firms and households produce differentiated goods and labor services, respectively. Only fractions $(1 - \xi_p)$ of firms and $(1 - \xi_w)$ of households are allowed to reoptimize their prices/wages in a given period. The likelihood of being selected for reoptimization is constant over time and independent across agents. Those that do not reoptimize fully index to past inflation. (M-10) and (M-11) are the linearized first-order conditions describing the share of firms and households that reoptimize in the current period.

Eq. (M-12) is the aggregate resource constraint. The parameter $\lambda_f \geq 1$ determines producers' average markup of price over marginal cost. It appears in (M-12) because firms encounter fixed costs that are assumed to be a constant share of output. Eq. (M-13) is a financial market-clearing condition requiring the demand for loans by firms equal the supply of deposits from households plus lump-sum transfers from the central bank. Eq. (M-14) linearizes the money growth rate. Finally, (M-15) is the monetary policy rule. It corresponds to the moving average representation for money growth implicit in a VAR (described below) in which policy shocks are identified as orthogonal innovations to the federal funds rate.

Because the VAR abstracts from all other non-monetary shocks, θ_j measures the response of money growth j periods from now to a policy innovation in the current period.

3 The Econometric Procedure

The goal of this paper is to see how fragile estimates of the Christiano-Eichenbaum-Evans model are to macroeconomic data revisions. To that end, I employ a strategy similar to one put forth by Croushore and Stark (2003) and Croushore and Evans (2006) in which the model parameters are estimated many times over using different vintages of historical data. Within each estimation attempt, I keep the sample period fixed in order to isolate the role of data revisions on the observed variability of the point estimates across vintage dates.

The model is estimated using the same limited information procedure adopted by the original authors. For each data vintage, a subset of the parameters listed in Table 2 is estimated by minimizing a weighted discrepancy between the model-based impulse response functions and their empirical counterparts taken from a recursively identified VAR.

3.1 A VAR Representation of the Data

Empirical response functions are derived from a covariance-stationary VAR

$$Y_t = \phi + A(L)Y_{t-1} + C\varepsilon_t,$$

where C is a square matrix and ε_t is a zero mean, serially uncorrelated vector of fundamental shocks with $E(\varepsilon_t\varepsilon_t') = I$. The variables comprising Y_t are partitioned into three groups. The group ordered first includes real output, real consumption, an aggregate price index, real investment, the real wage, and labor productivity. The second group contains only the federal funds rate. The last group consists of real profits and the growth rate of M2.⁶

Monetary shocks are characterized as orthogonal innovations to the funds rate. Under the above ordering, these correspond to realizations of the seventh fundamental shock in ε_t . Calculating the response of Y_t requires estimates of $A(L)$ and the seventh column of C , call it C_7 . I achieve identification by imposing $C_7 = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ c_7 \ c_8 \ c_9]'$. The six zero restrictions imply that monetary shocks have no contemporaneous effect on variables ordered above the funds rate. Leaving the last three coefficients free, however, allows the same shocks to have

⁶All of these variables, except for the federal funds rate and the growth rate of M2, have been logged.

an immediate impact on profits and the growth rate of M2.⁷

3.2 Parameter Estimation

Define $\boldsymbol{\gamma} \equiv (\lambda_f, \xi_w, \xi_p, \sigma_q, \kappa, b, \sigma_a)$ as the collection of parameters that are to be estimated. All other parameters listed in Table 2 are set equal to the calibrated values reported in Christiano *et al.* (2005). Denote $\boldsymbol{\Psi}(\boldsymbol{\gamma})$ the vector-valued function that maps $\boldsymbol{\gamma}$ to the model-based impulse responses, and let $\hat{\boldsymbol{\Psi}}$ be the corresponding vector of VAR-based response functions for a given vintage. Formally, the estimator for $\boldsymbol{\gamma}$ is

$$\hat{\boldsymbol{\gamma}} = \arg \min [\hat{\boldsymbol{\Psi}} - \boldsymbol{\Psi}(\boldsymbol{\gamma})]' \mathbf{V}^{-1} [\hat{\boldsymbol{\Psi}} - \boldsymbol{\Psi}(\boldsymbol{\gamma})],$$

where \mathbf{V} is a diagonal matrix containing the sample variances of each element in $\hat{\boldsymbol{\Psi}}$ along the main diagonal.⁸ In selecting \mathbf{V}^{-1} as the weight matrix, $\hat{\boldsymbol{\gamma}}$ places a higher priority on matching the empirical impulse responses that are more precisely estimated by the VAR.

3.3 The Real-Time Data Set

Estimates of $\hat{\boldsymbol{\Psi}}$ are obtained using a “real-time” data set that includes historical values for all of the variables represented in the VAR as they appeared in the second month of every quarter from 1997 through 2015.⁹ Each of the 76 consecutive vintages housed in the data set has a common sample period that runs from 1965:Q3 to 1996:Q3. The first quarter of 1997 was chosen as the initial vintage period because much of the data used in this paper, particularly data from the National Income and Product Accounts (NIPA), are first released in the quarters immediately following the activity date. Observations on all of the requisite

⁷Note that the model’s timing protocols are consistent with the identifying assumptions of the VAR. Date- t decisions for prices, wages, consumption, investment, and utilization are made before observing the policy shock. This is why expectations in (M-1)-(M-4) and (M-10)-(M-11) are all conditioned on information available through date $t - 1$. Cash balance and factor input decisions, however, are made after the shock occurs, allowing profits and money growth to respond contemporaneously rather than with a one-period lag.

⁸The diagonals of \mathbf{V} are obtained using Monte Carlo methods. I first take the joint distribution of the reduced-form VAR coefficients and the residual covariance matrix to be asymptotically normal with mean equaling the sample estimates and covariance equaling the sample covariance matrix of those estimates. I then draw 10,000 random vectors from this normal distribution and, preserving the identification restrictions, compute impulse response functions for each draw. Sample variances are calculated over the 10,000 trials.

⁹Each vintage contains the latest time series that would have been available to researchers on the 15th day of the second month of the quarter. For example, the vintage for the second quarter of 2006 contains sample data as it appeared in published volumes on May 15, 2006.

Table 3
Raw data series

Variable	Units	Frequency	Source	Mnemonic
Nominal GDP	Bil. \$, SAAR	Quarterly	RTDSM	NOUTPUT
Real GDP	Bil. real \$, SAAR	Quarterly	RTDSM	ROUTPUT
PCE: nondurable goods	Bil. \$, SAAR	Quarterly	ALFRED	PCND
PCE: services	Bil. \$, SAAR	Quarterly	ALFRED	PCESV
PCE: durable goods	Bil. \$, SAAR	Monthly	ALFRED	PCEDG
Gross Private Domestic Investment	Bil. \$, SAAR	Quarterly	ALFRED	GPDI
Government Cons. Exp. and Gross Inv.	Bil. \$, SAAR	Quarterly	ALFRED	GCE
Compensation Per Hour (nonfarm bus. sector)	Index, SA	Quarterly	ALFRED	COMPNFB
Output Per Hour All Persons (nonfarm bus. sector)	Index, SA	Quarterly	ALFRED	OPHNFB
Effective Federal Funds Rate	Annual %	Monthly	ALFRED	FEDFUNDS
Corp. Profits after Tax w/ IV Adj. and CC Adj.	Bil. \$, SAAR	Quarterly	RTDSM	NCPROFATW
M2 Money Stock	Bil. \$, SA	Monthly	RTDSM	M2

Notes: SA - seasonally adjusted; SAAR - seasonally adjusted annual rate; RTDSM - Real-Time Data Set for Macroeconomists; ALFRED - Archival Federal Reserve Economic Data. Monthly data were converted to quarterly by taking 3-month averages.

data series through 1996:Q3 were not actually available until early 1997.¹⁰

Table 3 provides details on all of the raw data series as well as the online sources from which the information was extracted. In some cases data were taken from the Real-Time Data Set for Macroeconomists (RTDSM) discussed in Croushore and Stark (2001) and available on the website of the Federal Reserve Bank of Philadelphia.¹¹ In cases where RTDSM did not contain the needed series, data were taken from Archival Federal Reserve Economic Data (ALFRED), a vast online database managed by the Federal Reserve Bank of St. Louis.¹²

The nine variables that characterize the VAR are assembled from the raw data using conceptual definitions outlined in Altig *et al.* (2011). Output is measured by real GDP (ROUTPUT). The aggregate price index corresponds to the ratio of nominal GDP (NOUTPUT) to real GDP. Nominal consumption is the sum of personal consumption expenditures on nondurable goods (PCND) and services (PCESV) as well as government consumption expenditures and gross investment (GCE). Nominal investment is the sum of personal consumption expenditures on durable goods (PCEDG) and gross private domestic investment

¹⁰Advance data on corporate profits, one of the NIPA variables used in estimation, is not released until the end of the second (or sometimes third) month of the quarter immediately following the activity date. Advance estimates of the other relevant NIPA variables are released at the end of the first month of the quarter following the activity date. As a result, the February 1997 vintage is the earliest vintage that contains observations on the full set of variables through 1996:Q3.

¹¹<http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/>

¹²<http://alfred.stlouisfed.org/>

(GPDI). Nominal wages are measured by compensation per hour in the non-farm business sector (COMPNFB). Nominal profits are after-tax corporate profits with inventory valuation and capital consumption adjustments (NCPROFATW). Consumption, investment, wages, and profits are converted to real terms by dividing by the aggregate price index. The interest rate is the effective federal funds rate (FEDFUNDS). Labor productivity is measured by real output per hour of all persons in the non-farm business sector (OPHNFB), and the money growth rate is the annualized growth rate of M2.

4 An Overview of the Revisions

This paper is principally concerned with measuring the impact of structural data revisions on DSGE coefficient estimates. These type of revisions, while infrequent, typically involve broad definitional changes that reshape the complete time series for affected variables. Before investigating their effects, however, it is important to get an idea of when these revisions occurred and what kinds of changes they brought to the data. Such information may be useful for putting into context observed shifts in the coefficient estimates over time.

One way to identify the past schedule of revisions is by tracking a given data point from the sample across all 76 vintages of the real-time data set. To that end, Fig. 1 shows year-over-year changes in the log of real investment *for the fourth quarter of 1995* as it appeared in consecutive quarters from 1997 through 2015.¹³ Recall that in the data investment is the sum of durable consumption and gross private domestic investment. Both measures are taken from the NIPA and assembled by the Bureau of Economic Analysis (BEA). It so happens that every four to five years the BEA makes major changes to the NIPA in what is termed a “comprehensive” revision. The timetable for these revisions can be determined by locating the vintages in which readings on investment suddenly change. Of course any series built exclusively from NIPA data could be used for the same purpose (e.g., real output, consumption, profits, and the price level). For sake of clarity, I only demonstrate the revisions to investment, leaving a discussion of the other NIPA concepts to the online appendix.¹⁴

In February 1997 the year-over-year growth rate of investment (for the fourth quarter of

¹³I look at year-over-year *changes* instead of log *levels* because the underlying data have been re-scaled over time as a result of base year adjustments. Such re-scaling, to which log-differencing is invariant, obscures the amount by which individual data points get revised from one vintage to the next.

¹⁴The appendix is available from the journal website at <http://www.jmcb.osu.edu/>. There I construct a revision “matrix” that identifies, series-by-series, every vintage containing a structural or information-based revision to at least one data point in the sample. Accompanying this matrix is a detailed discussion of all the structural revisions that affected each series listed in Table 3.

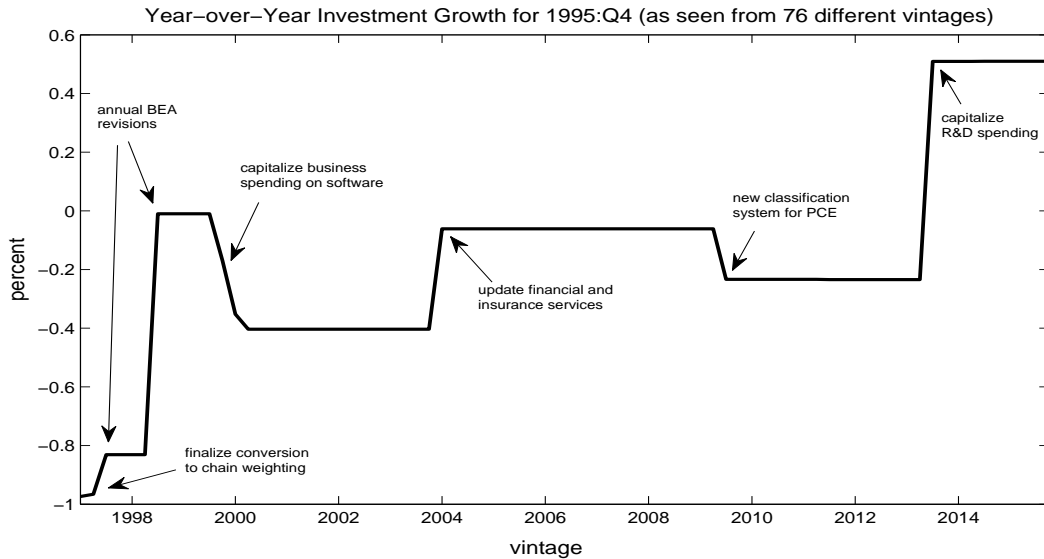


Fig. 1. Readings on the year-over-year growth rate of real investment for the fourth quarter of 1995 are shown for vintages dated February 1997 to November 2015. In the data investment is durable consumption plus gross private domestic investment.

1995) stood at -1 percent but was revised up to -0.8 percent by August of the same year. Just twelve months later it was revised up again to 0 percent. The estimate was subsequently pushed back down to -0.4 percent in a series of revisions that took place between November 1999 and May 2000. The growth rate then stayed at this level until February 2004, at which point it got revised up to around -0.1 percent. The next changes came in August of 2009 and 2013. The 2009 revision witnessed investment growth drop to -0.2 percent, but the 2013 revision saw it rise sharply into positive territory, or 0.5 percent to be exact.

As expected, most of the changes to investment coincide with the release of a comprehensive revision to the NIPA. Over the last twenty years (as of this writing), five such revisions have occurred: in January 1996, October 1999, December 2003, July 2009, and July 2013. The 1996 revision marked the beginning of the government's switch from fixed-weight to chain-weight methods for calculating real quantities. Changes to all of the historical data, however, were not completed until May 1997, or one quarter after the initial vintage date for this study. In the October 1999 revision, the BEA reclassified software expenditures by businesses and government as fixed investment; it was previously regarded as an office expense. The full set of revisions to durable consumption and real GDP, however, were not finished until December and early 2000, respectively. This explains why real investment for the fourth quarter of 1995 continued to change through the first half of 2000.

The next comprehensive revision came in December 2003 and incorporated, among other

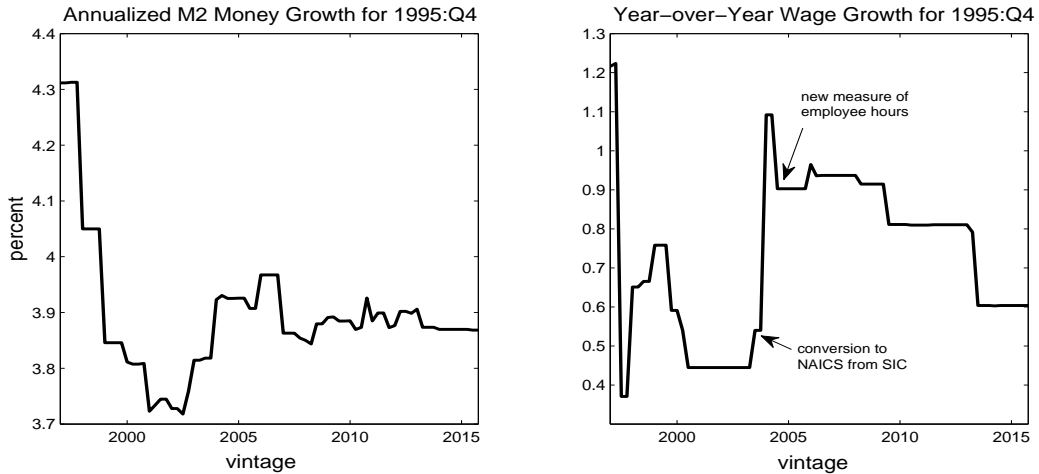


Fig. 2. Readings on the growth rate of M2 and the year-over-year growth rate of the real wage for the fourth quarter of 1995 are plotted for vintages dated February 1997 to November 2015. In the data wages correspond to nonfarm hourly compensation.

things, improved measures of services administered by insurance providers, banks, and government agencies. The July 2009 revision featured two big changes to the NIPA. It updated the accounting of disaster-related damages to fixed assets, and it modernized the entire classification structure for personal consumption expenditures. The last comprehensive revision in July 2013 centered on the creation of a new category of private fixed investment—intellectual property products—that lumped together spending by businesses and nonprofits on things like research and development and artistic originals. These were formerly counted as intermediate inputs rather than as investments in a durable asset.

In contrast to investment, some of the data used to estimate the Christiano-Eichenbaum-Evans model often get revised for reasons unrelated to any comprehensive change in the NIPA. This group includes the real wage, labor productivity, and M2 money growth. The first two are assembled using data from the Bureau of Labor Statistics (BLS) on compensation and output per hour in the nonfarm business sector. Data on M2 is published by the Federal Reserve Board. The chronology of past revisions can once again be verified by following a single data point across vintage time. Fig. 2 plots the growth rate of M2 (left panel) and the year-over-year growth rate of the real wage (right panel) for the fourth quarter of 1995 as it appeared in real time from 1997 through 2015. Not shown in the figure is a vintage plot of labor productivity. As discussed in the online appendix, revisions to this series are almost perfectly synchronized with revisions to hourly compensation.

It is clear that revisions to M2 are much more frequent than revisions to investment or any other NIPA concept for that matter. Between 1997 and 2015, the estimate of annualized M2

growth for 1995:Q4 changed 45 times (out of 76 vintages). The majority of these revisions, however, were small, averaging about three basis points up or down. According to reports issued by the Federal Reserve, M2 revisions are almost always caused by changes to the source data supplied by depository institutions or changes in seasonal factors. Since neither involve major definitional changes to the money supply, they should probably be viewed as information-based rather than structural.

Revisions to the real wage also appear to be relatively frequent. Since 1997, the estimate of real wage growth for the fourth quarter of 1995 changed on 23 separate occasions. The average size of one of these revisions was about twelve basis points. Unlike M2, however, some revisions have been structural while others have been strictly information-based. As detailed in the online appendix, non-farm hourly compensation is constructed using wage and salary data from the NIPA along with hours data from the BLS Current Employment Statistics (CES) program and the Census Bureau's Current Population Survey (CPS). Any definitional change to one or more of these primary sources would thus result in a structural revision to the wage series. The comprehensive revisions to the NIPA discussed earlier certainly fit this description, but they are not the only ones. In July 2003, for example, CES employment and hours data were revised following the government's conversion from the Standard Industrial Classification (SIC) to the North American Industry Classification System (NAICS). A year later, new measures of employee hours for all of the major business sectors were developed from data in the CPS. Both sets of revisions contributed to the large swing in the estimate of non-farm hourly compensation observed between 2003 and 2004.

4.1 Discussion

Putting aside the details, it should come as no surprise that large-scale revisions to the data like the ones described above are necessary in an ever-changing economy. As new products enter the market and as spending patterns evolve, national accounting frameworks must be updated to accommodate these changes. Unfortunately, a majority of theoretical models, the present one included, have nothing to say about which accounting methods are relevant for structural estimation. For example, should analysts use investment data that accounts for expenditures on software, research and development, and other intellectual property products? Should real quantities be evaluated using fixed-weight or chain-weight methods? These are questions about which most DSGE models provide no clear-cut answers.

So with little guidance from theory, how should an empirical researcher decide which vintage of data to use? Croushore and Stark (2003) contend that in fact such a decision

should not be made at all. Instead, models should be estimated on several different vintages, both past and present, in an effort to demonstrate robustness. Here the authors make a simple yet powerful argument. For the empirical validity of a model to be fully trusted, the estimation results ought to hold up under changes in the sample data about which the model itself is largely ambiguous. For DSGE models this means the coefficient estimates should be fairly stable or robust across vintages that employ different accounting standards but describe the same basic economic concepts. A lack of robustness, on the other hand, would not only raise doubts over the coherence between one’s model and the data, it would call into question its very usefulness as a tool for interpreting observed macroeconomic behavior.

5 Do Revisions Affect Estimation Results?

This section examines how data revisions affect estimates of the Christiano-Eichenbaum-Evans model. The central findings are illustrated in Fig. 3, which plots sequential estimates of γ for quarterly vintages spanning 1997 through 2015. In each panel the solid line shows the point estimates (left scale) and the dotted line shows the corresponding standard errors (right scale).¹⁵ Table 4 provides some descriptive statistics summarizing these estimates.

5.1 Parameter Estimates

Fig. 3 indicates that estimates of the model are generally robust to the historical data revisions that occurred between 1997 and 2015. This is not to say that they are invariant across vintages. On the contrary, all seven parameter estimates undergo continual changes, and several of the most prominent happen on or near dates associated with the big structural revisions discussed in the previous section. By robust, I simply mean that the observed variations in $\hat{\gamma}$ over time are relatively small compared to the *sampling* uncertainty surrounding the point estimate from any given vintage.

Consider, for example, the price markup coefficient λ_f . The mean point estimate over all 76 vintages is 1.20, the same as the value reported in Christiano *et al.* (2005). Around the mean I find that point estimates vary between 1.11 and 1.39 with a standard deviation of 0.08 (see Table 4). Nevertheless, what stands out is that the variability in these estimates across vintages is smaller than the typical sampling variability of $\hat{\lambda}_f$ as measured by its standard error. Note that the standard error estimates average 0.29 and are uniformly greater than

¹⁵Standard errors are computed using the asymptotic delta method as described in Christiano *et al.* (2001).

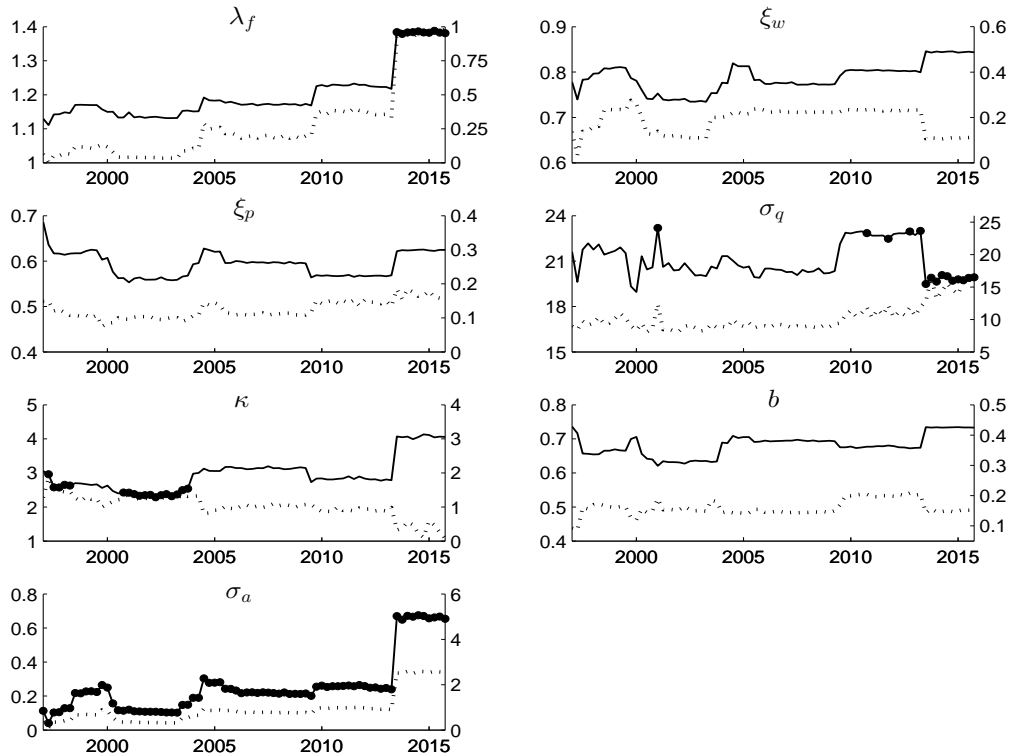


Fig. 3. Parameter estimates are graphed for vintages dated February 1997 to November 2015. Solid lines mark the point estimates (left scale) and dotted lines mark the standard errors (right scale). Bullets indicate vintages in which, using a standard Wald test, the null hypothesis that the true parameter value equals zero cannot be rejected at a five percent significance level.

0.10 for vintages published after the comprehensive NIPA revision in December 2003.

A similar pattern emerges for the other coefficients. Take, for instance, the Calvo probabilities ξ_w and ξ_p , which pin down the average length of wage and price contracts in the economy. Point estimates of the former fluctuate between 0.73 and 0.85, while estimates of the latter vary from 0.55 to 0.69. The standard deviation computed over all vintages is only around 0.03 for both parameters. By contrast, the standard errors associated with vintage estimates of ξ_w and ξ_p are usually much higher, averaging 0.19 and 0.12, respectively. Turning to the habit formation parameter, one sees that the variability in \hat{b} over time is small relative to its estimated sampling variability. Estimates of b , which average 0.68, have a standard deviation roughly five times smaller than the mean standard error estimate of 0.16.

Table 4
Descriptive statistics for vintage estimates

parameter	mean	min	max	std dev	mean(s.e.)	std dev(s.e.)	Wald
λ_f	1.201	1.111	1.388	0.078	0.283	0.288	87
ξ_w	0.789	0.734	0.846	0.032	0.187	0.056	100
ξ_p	0.593	0.553	0.686	0.027	0.124	0.026	100
σ_q	21.035	18.964	23.190	1.146	10.272	2.126	80
κ	2.964	2.287	4.130	0.507	1.017	0.331	76
b	0.681	0.621	0.736	0.032	0.160	0.025	100
σ_a	0.259	0.041	0.675	0.169	0.932	0.675	0

Notes: The table reports the mean, min, max, and standard deviation of the point estimates for each parameter evaluated over all data vintages from February 1997 to November 2015. It also reports the mean and standard deviation of the estimated standard errors (s.e.) associated with each parameter. The last column reports the percentage of all samples in which the parameter estimate is statistically different from zero (5% significance level) according to a standard Wald test.

5.2 Statistical Inference

The evidence presented thus far suggests that estimates of the model parameters are themselves not overly sensitive to variations in the sample caused by data revisions. Where revisions appear to have more influence is on the statistical (in)significance of $\hat{\gamma}$, that is, on hypothesis tests about the true value of γ . The reason is that for at least some parameters, the standard error estimates change a lot following major revisions to the data. Because many common testing procedures rely on standard errors, abrupt changes in their values can have a big effect on the outcome of hypothesis testing.

An obvious example of this concerns estimates of σ_q , the (inverse) interest semielasticity of money demand. For vintages that predate the comprehensive NIPA revision of July 2013, a classical Wald test of the null hypothesis that $\sigma_q = 0$ is rejected at a five percent significance level in 61 of 66 cases.¹⁶ For vintages published after this date, however, the number of test rejections is zero. Fig. 3 makes clear that what drives this result is a shift up in the estimated standard errors. Between 2009 and 2013, the standard errors center closely around 11. After July 2013 the typical standard error jumps to about 15, rendering the point estimates statistically insignificant. Overall, $\hat{\sigma}_q$ is significantly different from zero in 80 percent of vintages comprising the real-time data set (see Table 4).

Inferences about the true value of λ_f are also not fully robust. Once again, the standard error estimates rise dramatically upon release of the July 2013 NIPA revision, surging from

¹⁶The Wald statistic is formed by squaring the ratio of the point estimate to its standard error. Under the null hypothesis, it is asymptotically distributed chi-square with one degree of freedom.

a plateau of about 0.37 to a higher level of 0.96. The main consequence of this change is that point estimates of λ_f are statistically significant in all vintages prior to the revision but uniformly insignificant thereafter.

Taken together, the preceding results suggest that analysts should exercise caution when assessing the coefficient estimates of popular DSGE models using only a single vintage of macroeconomic data. For doing so leaves open the question of whether the findings are robust against alternative vintages. In the case of the Christiano-Eichenbaum-Evans model, questions about robustness concern not so much the values of the parameter estimates per se, but rather the outcome of hypothesis tests. For instance, research conducted in 1997 would have concluded, with high confidence, that estimates of λ_f and σ_q are large and statistically significant. Yet, the exact same analysis performed in 2015 would have found scant evidence of statistical significance. The conflict here is plainly a consequence of the revisions that took place during the intervening years—the 2013 comprehensive NIPA revision in particular—which apparently exposed $\hat{\lambda}_f$ and $\hat{\sigma}_q$ to additional sampling uncertainty.

It is worth pointing out that data revisions have in some cases improved the precision of statistical inference. Nowhere is this more evident than in estimates of κ , the (inverse) price elasticity of investment. In vintages published before the December 2003 comprehensive revision, a Wald test of the hypothesis that $\kappa = 0$ is rejected (at a 5 percent level) only 10 times out of 28. In subsequent vintages, the rejection rate is 100 percent. Accompanying this increase in parameter significance are sizable reductions in the standard errors. Estimates average about 1.30 prior to February 2004, 0.98 from that point until August 2013, and 0.34 thereafter. So unlike λ_f and σ_q , estimates of κ have actually become more precise over time.

5.3 Which Revisions Matter Most?

The results depicted in Fig. 3 suggest that certain data revisions may have had a bigger effect on estimation and inference than others. To identify which particular revisions matter most for DSGE estimation, I first difference the sequence of vintage estimates, $\{\hat{\gamma}\}$, and then take the absolute value of the resulting series. This transformation reveals by how much the point estimates change from one vintage to the next. In Table 5 I list the values (in parentheses) and dates of the three biggest parameter shifts in descending order.

Of the major structural revisions that occurred between 1997 and 2015, one clearly stands out as having the most impact on estimates of the Christiano-Eichenbaum-Evans model, the July 2013 comprehensive revision to the NIPA. For all seven parameters, the break in the point estimate observed in the August 2013 vintage is larger than that of any other vintage.

Table 5
Vintages producing the largest breaks in parameter estimates

parameter	1		2		3	
λ_f	Aug. 2013	(0.167)	Nov. 2009	(0.057)	Aug. 2004	(0.040)
ξ_w	Aug. 2013	(0.046)	Aug. 2004	(0.044)	Aug. 1997	(0.043)
ξ_p	Aug. 2013	(0.053)	May 1997	(0.050)	Aug. 2004	(0.032)
σ_q	Aug. 2013	(3.502)	Feb. 2001	(2.607)	May 2001	(2.571)
κ	Aug. 2013	(1.285)	Feb. 2004	(0.440)	Aug. 2009	(0.414)
b	Aug. 2013	(0.060)	Aug. 1997	(0.059)	Feb. 2004	(0.055)
σ_a	Aug. 2013	(0.431)	Aug. 2004	(0.114)	May 2000	(0.093)

Notes: The table reports the three vintages for which the parameter estimates change (from the previous vintage) by the largest absolute amount. Numbers in parentheses identify the size of the parameter shift and are listed in descending order.

This is especially true of λ_f , κ , and σ_a , where the break is three to four times the size of its closest rival. The estimate of σ_a , which measures the (inverse) elasticity of capital utilization, actually jumps by 180 percent compared to its May 2013 value. Whether this observation is significant, however, is unclear since $\hat{\sigma}_a$ is never statistically different from zero.¹⁷

One change that does appear significant concerns the estimate of κ . Upon release of the July 2013 revision, $\hat{\kappa}$ jumps from 2.79 to 4.07 while its standard error falls by about two-thirds. So not only is it greater than zero, $\hat{\kappa}$ is also statistically different from its previous estimate of 2.79.¹⁸ Of course this finding raises an obvious question. Is there something unique about the 2013 comprehensive revision that would affect estimates of κ over and above the other parameters? As noted in section 4 and explained in detail by McCulla, Holdren, and Smith (2013), several improvements to the NIPA were implemented on July 31 of that year. Chief among them was a reclassification of spending by private enterprises on research and development as well as the creation of entertainment, literary, and artistic originals as fixed investment rather than intermediate inputs. These wholesale changes to the way investment was to be measured substantially altered historical readings on many important NIPA variables, the most prominent being gross private domestic investment (see Fig. 1). Because GPDI is a key component of the investment series used for estimation, large-scale revisions have the potential to affect inferences about the structural model, particularly those features that regulate investment dynamics. Thus it is not unreasonable to expect that the 2013 revision would have a discernible effect on both the size and precision of $\hat{\kappa}$. Recall

¹⁷Christiano *et al.* (2005) report that $\hat{\sigma}_a$ converges to zero during the course of estimation. As a result, they fix $\sigma_a = 0.01$ and optimize over the remaining elements of γ .

¹⁸Using the August 2013 vintage, the p -value for a Wald test of the null that $\kappa = 2.79$ is less than 0.01.

that $1/\kappa$ is the elasticity of investment with respect to a unit increase in the price of capital, and as such, plays a central role in shaping the behavior of investment activity in the model.

6 Effects of Variations in Sample Size

The estimation strategy employed in the last section, whereby the sample period is held fixed across vintages, was designed to draw maximum attention to the broad-based structural data revisions that occurred between 1997 and 2015. It is after all the effect these revisions have on parameter estimates that is the central focus of this study. Nevertheless, there are a couple of reasons why it may be worthwhile to see what happens when the sample period gets updated over time. For starters, demonstrating robustness to additional data, rather than data revisions alone, would bolster the argument that estimates of the Christiano-Eichenbaum-Evans model stand the test of time. A second reason is that it would make for a better comparison with results found in Tetlow and Ironside (2007) and Tetlow (2012), both of whom analyze different vintages of the FRB/US model.

Table 6 reports estimates of the model from six vintages of data. Included are the first and last vintages in the real-time data set (February 1997 and November 2015) along with vintages immediately preceding the last four comprehensive NIPA revisions (August 1999, November 2003, May 2009, and May 2013). Now in contrast to the benchmark analysis, the sample from each vintage contains the latest observations. So the November 2003 vintage, for instance, has a sample period covering 1965:Q3 to 2003:Q2 instead of 1965:Q3 to 1996:Q3. It follows that changes in the estimates over time (moving down the table) are affected not just by the occurrence of revisions, but also by the addition of new sample data.

Estimates of the Christiano-Eichenbaum-Evans model do not appear as robust to changes in the sample size as they are to data revisions alone. Some parameters get noticeably larger as more and more data are added to later vintages. Obvious examples here are the markup coefficient λ_f , the Calvo probability ξ_p , and the habit factor b . In each case, the estimate obtained using data through 2015:Q2 is statistically different from the estimate implied by the February 1997 vintage according to a Wald test.¹⁹ Values for σ_a also get bigger over time, so much so in fact that the implied elasticity of capital utilization ($1/\sigma_a$) swiftly approaches zero. Vintage estimates of σ_a , however, are all highly imprecise, and as a result, have less bearing on the model's overall ability to fit the impulse response functions. Estimates of ξ_w ,

¹⁹Based on the November 2015 vintage, separate Wald tests of the hypotheses that λ_f , ξ_p , and b equal their February 1997 vintage estimates are rejected at one percent significance levels.

Table 6
Variations in sample size

vintage	sample	λ_f	ξ_w	ξ_p	σ_q	κ	b	σ_a
Feb. 1997	1965:Q3 to 1996:Q3	1.13 (0.06)	0.78 (0.14)	0.69 (0.15)	21.63 (8.94)	3.06 (1.28)	0.74 (0.09)	0.11 (0.37)
Aug. 1999	1965:Q3 to 1999:Q1	1.37 (0.92)	0.96 (0.15)	0.78 (0.30)	21.49 (13.21)	3.03 (2.03)	0.78 (0.13)	0.54 (1.53)
Nov. 2003	1965:Q3 to 2003:Q2	1.26 (0.29)	0.90 (0.16)	0.64 (0.14)	20.97 (5.84)	2.72 (0.26)	0.75 (0.10)	0.35 (0.58)
May 2009	1965:Q3 to 2008:Q4	2.51 (0.63)	0.38 (0.32)	0.99 (0.04)	22.99 (11.14)	2.01 (0.86)	0.75 (0.03)	1000 [†]
May 2013	1965:Q3 to 2012:Q4	1.64 (0.17)	0.80 (0.26)	0.99 (0.04)	21.16 (12.18)	2.26 (0.97)	0.83 (0.01)	1000 [†]
Nov. 2015	1965:Q3 to 2015:Q2	1.70 (0.07)	0.80 (0.22)	0.99 (0.04)	20.75 (5.66)	2.45 (0.77)	0.84 (0.01)	1000 [†]

Notes: The table reports estimates and standard errors (in parentheses) for select vintages. In each case the sample period is updated to include the most recent observations available at the time. † denotes a value that is imposed prior to estimation.

σ_q , and κ , on the other hand, are relatively more stable across vintages. In fact, none of their November 2015 estimates are statistically different from the initial February 1997 estimates.

The results in Table 6 make clear that lengthening the sample period substantially alters estimates of certain parameters. Yet it remains to be seen whether, or to what extent, these changes affect the main quantitative properties of the model. To that end, I follow Tetlow and Ironside (2007) and Tetlow (2012) in computing the *real-time* policy multiplier implied by the estimated model. The multiplier is defined here as the cumulative loss of real output after eight quarters that results from an unanticipated 50-basis-point increase in the federal funds rate.²⁰ Fig. 4 graphs the real-time multiplier by vintage for the period 1997 to 2015. For each quarterly vintage, two versions of the multiplier are displayed. Points on the solid line show the cumulative output response produced by the model vintage of that date and estimated using a sample that includes the most recent observations available at the time. Points on the dashed line show the outcome of the same policy experiment, but with each model vintage estimated over the benchmark sample period of 1965:Q3 to 1996:Q3. Comparing the two lines helps separate the joint effects that *revisions* and *additions* to the data have on real-time estimates of the policy multiplier.

Consider first the multipliers associated with fixed-sample estimates of the Christiano-Eichenbaum-Evans model. For the most part, estimates of the output response in real time are fairly stable over the period in question. Indeed, from February 1997 to May 2013 the

²⁰Monetary policy is represented in the model by (M-15), an infinite-order moving average process for the growth rate of money. In computing the multiplier, I calibrate the money growth innovation so that the model delivers a contemporaneous increase in the short-term nominal interest rate of exactly 50 basis points.

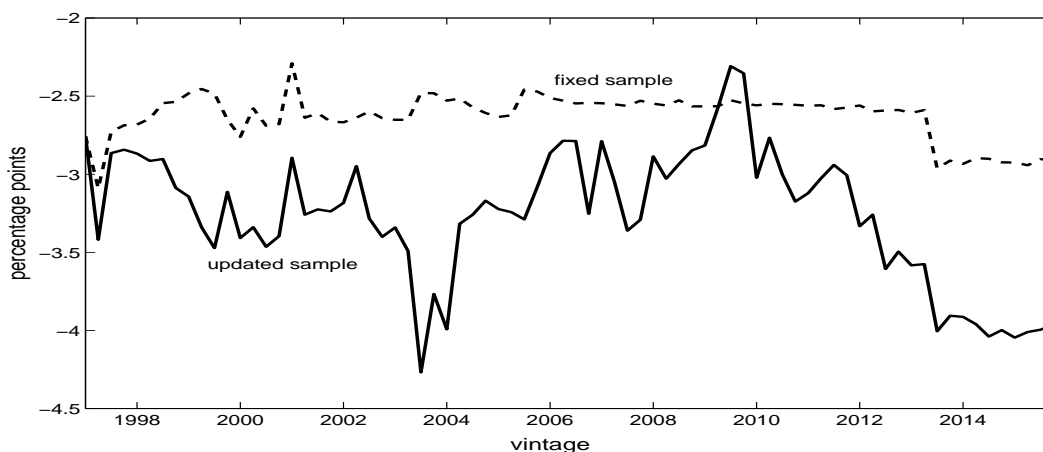


Fig. 4. Real-time policy multipliers implied by the estimated model are graphed by vintage from February 1997 to November 2015. The dashed line comes from estimates obtained using the benchmark sample period 1965:Q3 to 1996:Q3. The solid line is derived using progressively larger samples that contain the latest observations available at the time. The policy multiplier is defined as the cumulative two-year response of real output to an unexpected 50-basis-point rise in the federal funds rate.

model regularly predicted that output would fall by an average of about 2.6 percent (over two years) following a surprise 50-basis-point increase in the funds rate. But after the July 2013 revision to the NIPA, the forecast jumped to around -2.9 percent and stayed at that level through 2015. The timing of this change is not surprising given results in the previous section which point to the August 2013 vintage as having a significant impact on estimates of key structural parameters. The other revisions tend to have a smaller effect on model estimates and therefore less influence on the assessment of policy multipliers in real time.

Turning to the solid line, one sees that the multiplier effects look different when the sample period gets updated every quarter. Not only are the responses more volatile across vintages, they are almost always bigger (in an absolute sense) than their fixed-sample counterparts. From 1997 through 2015, estimates of the policy multiplier averaged -3.3 percent, but it also varied between -4.3 and -2.3 percent along the way. Interestingly, these results are comparable to ones found by Tetlow (2012) and Tetlow and Ironside (2007). In the context of FRB/US, the authors report similar quantities for both the magnitude and volatility of the funds rate multiplier from 1996 to 2007. The volatility in particular demonstrates that the real effects of monetary policy as recognized today can differ substantially from what analysts would have thought at certain points in the past. Like the present study, uncertainty around the multipliers traces directly to changes over time in estimates of the structural model. What the comparisons in Fig. 4 reveal is that in the case of the Christiano-Eichenbaum-Evans model, those changes result primarily from the use of more data, not revised data.

7 Concluding Remarks

Put simply, DSGE methodology is nothing more than a framework for interpreting macroeconomic data. What distinguishes it from its predecessors is the explicit use of coherent microeconomic foundations. But like any other framework, for the interpretations born out of DSGE to be regarded as empirically valid, the estimation results should hold up to data revisions, particularly the broad structural revisions about which most theoretical models are ambiguous. If instead the results are highly sensitive to changes in the data, then the model may yield little insight on the sources of key business cycle relationships.

This paper studies the impact of data revisions on estimates of the celebrated Christiano-Eichenbaum-Evans model of the US economy. The effects are identified by re-estimating the model on successive vintages of macroeconomic data while keeping the sample period fixed throughout. The results show that where parameter estimates are generally stable across vintages, standard error estimates are by comparison relatively more volatile. For some parameters, the break in the standard error observed in certain quarters is big enough to alter the outcome of a Wald test that its true value equals zero. So even though actual point estimates of the model coefficients appear robust to data revisions, judgements about their statistical significance can differ depending on which vintage is used for estimation.

The results described above speak to the robustness of the estimates of a particular DSGE model. Whether similar findings emerge using alternative models is an open question that warrants further investigation. At present, there are many well-known DSGE models emphasizing important aspects of the economy that could benefit from exposure to the kind of replication experiments carried out in this paper. Some examples include models that feature housing and collateral (e.g., Iacoviello, 2005), labor search externalities (e.g., Gertler, Sala, and Trigari, 2008), and financial market frictions (e.g., Del Negro, Eusepi, Giannoni, Sbordone, Tambalotti, Cocci, Hasegawa, and Linder, 2013).

Finally, it may be worthwhile to incorporate data revisions directly into the information structure of DSGE models. Croushore (2011) points out that both private agents and policymakers in the real world have to filter noisy incoming data when forming optimal plans. The nature of this data and the revisions that follow should, in principle, affect their decision making and in the process reshape the equilibrium dynamics of the model. Although an extension like this one would undoubtedly complicate the estimation procedure, the literature does provide at least a few examples from the real business cycle tradition (e.g., Bomfim, 2001; Aruoba, 2004) that suggest such a task is feasible.

References

- Altig, David; Christiano, Lawrence J.; Eichenbaum, Martin and Lindé, Jesper.** “Firm-Specific Capital, Nominal Rigidities and the Business Cycle.” *Review of Economic Dynamics*, April 2011, 14(2), pp. 225-47.
- Aruoba, S. Boragan.** “Data Uncertainty in General Equilibrium.” Unpublished Manuscript, May 2004.
- Blanchard, Olivier Jean and Quah, Danny.** “The Dynamic Effects of Aggregate Demand and Supply Disturbances.” *The American Economic Review*, September 1989, 79(4), pp. 655-73.
- Bomfim, Antulio N.** “Measurement Error in General Equilibrium: the Aggregate Effects of Noisy Economic Indicators.” *Journal of Monetary Economics*, December 2001, 48(3), pp. 585-603.
- Calvo, Guillermo A.** “Staggered Prices in a Utility-Maximizing Framework.” *Journal of Monetary Economics*, September 1983, 12(3), pp. 383-98.
- Christiano, Lawrence J.; Eichenbaum, Martin and Evans, Charles L.** “The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds.” *The Review of Economics and Statistics*, February 1996, 78(1), pp. 16-34.
- _____. “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy.” NBER Working Paper No. 8403, July 2001.
- _____. “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy.” *Journal of Political Economy*, January 2005, 113(1), pp. 1-45.
- Croushore, Dean.** “Forecasting with Real-Time Macroeconomic Data.” in G. Elliott and C. Granger and A. Timmermann, eds., *Handbook of Economic Forecasting*. Amsterdam and Oxford: Elsevier, North-Holland, 2006, pp. 961-82.
- _____. “Frontiers of Real-Time Data Analysis.” *Journal of Economic Perspectives*, March 2011, 49(1), pp. 72-100.
- Croushore, Dean and Evans, Charles L.** “Data Revisions and the Identification of Monetary Policy Shocks.” *Journal of Monetary Economics*, September 2006, 53(6), pp. 1135-60.
- Croushore, Dean and Stark, Tom.** “A Real-Time Data Set for Macroeconomists.” *Journal of Econometrics*, November 2001, 105(1), pp. 111-30.

- _____. “A Real-Time Data Set for Macroeconomists: Does the Data Vintage Matter?” *The Review of Economics and Statistics*, August 2003, 85(3), pp. 605-17.
- Croushore, Dean and Sill, Keith.** “Analyzing Data Revisions with a Dynamic Stochastic General Equilibrium Model.” Federal Reserve Bank of Philadelphia Working Paper No. 14-29, September 2014.
- Del Negro, Marco; Eusepi, Stefano; Giannoni, Marc; Sbordone, Argia; Tambalotti, Andrea; Cocci, Matthew; Hasegawa, Raiden and Linder, M. Henry.** “The FRBNY DSGE Model.” Federal Reserve Bank of New York Staff Report No. 647, October 2013.
- Galí, Jordi.** “How Well Does the IS-LM Model Fit Postwar U.S. Data?” *The Quarterly Journal of Economics*, May 1992, 107(2), pp. 709-38.
- Gertler, Mark; Sala, Luca and Trigari, Antonella.** “An Estimated Monetary DSGE Model with Unemployment and Staggered Nominal Wage Bargaining.” *Journal of Money, Credit and Banking*, December 2008, 40(8), pp. 1713-64.
- Hall, Robert E.** “Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence.” *Journal of Political Economy*, December 1978, 86(6), pp. 971-87.
- Iacoviello, Matteo.** “House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle.” *The American Economic Review*, June 2005, 95(3), pp. 739-64.
- Kydland, Finn E. and Prescott, Edward C.** “Business Cycles: Real Facts and a Monetary Myth.” Federal Reserve Bank of Minneapolis *Quarterly Review*, Spring 1990, pp. 3-18.
- McCulla, Stephanie H.; Holdren, Alyssa E. and Smith, Shelly.** “Improved Estimates of the National Income and Product Accounts.” *Survey of Current Business*, September 2013, 93(9), pp. 14-45.
- Nikolsko-Rzhevskyy, Alex.** “Monetary Policy Estimation in Real Time: Forward-Looking Taylor Rules without Forward-Looking Data.” *Journal of Money, Credit and Banking*, August 2011, 43(5), pp. 871-97.
- Orphanides, Athanasios.** “Monetary Policy Rules Based on Real-Time Data.” *The American Economic Review*, September 2001, 91(4), pp. 964-85.
- Smets, Frank and Wouters, Raf.** “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach.” *The American Economic Review*, June 2007, 97(3), pp. 586-606.

- Swanson, Norman.** “Forecasting Using First-Available Versus Fully Revised Economic Time-Series Data.” *Studies in Nonlinear Dynamics and Econometrics*, April 1996, 1(1), pp. 47-64.
- Tetlow, Robert J.** “Real-Time Model Uncertainty in the United States: ‘Robust’ Policies Put to the Test.” *International Journal of Central Banking*, March 2015, 11(1), pp. 113-55.
- Tetlow, Robert J. and Ironside, Brian.** “Real-Time Model Uncertainty in the United States: The Fed, 1996-2003.” *Journal of Money, Credit and Banking*, October 2007, 39(7), pp. 1533-61.
- Yun, Tack.** “Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles.” *Journal of Monetary Economics*, April 1996, 37(2), pp. 345-70.