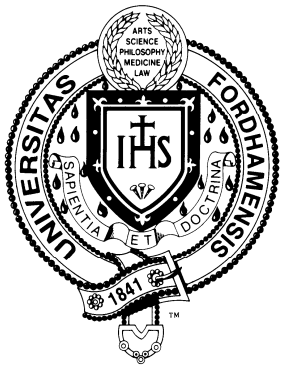

Fordham University
Department of Economics
Discussion Paper Series

GMM and OLS Estimation and Inference for New Keynesian Phillips Curve

Hrishikesh D. Vinod
Fordham University, Department of Economics

Discussion Paper No: 2010-02
March 2010

Department of Economics
Fordham University
441 E Fordham Rd, Dealy Hall
Bronx, NY 10458
(718) 817-4048

GMM and OLS Estimation and Inference for New Keynesian Phillips Curve

H. D. Vinod *

Abstract

This paper considers estimation situations where identification, endogeneity and non-spherical regression error problems are present. Instead of always using GMM despite weak instruments to solve the endogeneity, it is possible to first check whether endogeneity is serious enough to cause inconsistency in the particular problem at hand. We show how to use Maximum Entropy bootstrap (meboot) for non-stationary time series data and check 'convergence in probability' and 'almost sure convergence' by evaluating the proportion of sample paths straying outside error bounds as the sample size increases. The new Keynesian Phillips curve (NKPC) ordinary least squares (OLS) estimation for US data finds little endogeneity-induced inconsistency and that GMM seems to worsen it. The potential 'lack of identification' problem is solved by replacing the traditional pivot which divides an estimate by its standard error by the Godambe pivot, as explained in Vinod (2008) and Vinod (2010), leading to superior confidence intervals for deep parameters of the NKPC model.

*Professor of Economics, Fordham University, Bronx, New York, USA 10458. url: http://www.fordham.edu/economics/vinod E-Mail: vinod@fordham.edu, I thank Prof. Kleibergen for sending me the data for the NKPC model, Professors P. Chausse, P. Lafaye de Micheaux and B. Moore for helpful comments. I am responsible for any remaining errors.

KEY WORDS: bootstrap; simulation; convergence; inflation inertia; sticky prices.

1 Introduction

This paper considers inference issues in models where endogeneity of regressors is feared and non-spherical regression errors are present, in the context of an important macroeconomic model which has become a workhorse for analyzing monetary policy and fluctuations. Galí and Gertler (1999) published an important paper dealing with the micro foundations of the new Keynesian Phillips curve (NKPC) where some fraction ϕ of firms use a backward looking rule-of-thumb when they set their prices. Their work has inspired several papers cited in Kleibergen and Mavroeidis (2009), (hereafter "KM09") which is followed by several discussion papers in the same issue of the journal. KM09 focus on identification, testing and inference for NKPC within the generalized method of moments (GMM) framework.

The inflation π_t is measured by the deflator for quarterly US gross domestic product (GDP) as a part of Federal Reserve Economic Data (FRED) for (1947-2007) at `http://research.stlouisfed.org/fred2`. We use KM09 data for the regressor x_t as a proxy for real marginal costs based on the share of labor in the nonfarm business sector developed from the data published by the Bureau of Labor Statistics.

If E_t denotes expectation conditional on information at time t, the NKPC model is given by:

$$\pi_t = \lambda x_t + \gamma_f E_t(\pi_{t+1}) + \gamma_b \pi_{t-1} + u_t. \tag{1}$$

Economists are more interested in three deep parameters: β the discount factor, α the probability that prices remain fixed, and ϕ the fraction of backward looking price setters in the economy. The deep parameters are related to the parameters of (1) by the following nonlinear equations derived by Galí and Gertler (1999). They represent three equations in three unknowns, apart from a definitional identity for A.

$$
\begin{aligned}
A &= \alpha + \phi\,(1 - \alpha\,(1 - \beta)) \\
\lambda &= (1 - \phi)(1 - \alpha)(1 - \beta\,\alpha)/A \\
\gamma_f &= \beta\,\alpha/A \\
\gamma_b &= \phi/A
\end{aligned} \tag{2}
$$

We shall see that numerical estimates of the deep parameters β, α and ϕ can be obtained even though simple explicit expressions are unavailable. We simply use the R Development Core Team (2008) package 'rootSolve' and a suitable set of estimates of parameters $\{\lambda, \gamma_f, \gamma_b\}$ for the left sides.

It is customary to replace the $E_t(\pi_{t+1})$ in eq. (1) by $\pi_{t+1} - \eta_{t+1}$, where η_t denotes one-step ahead forecast error for inflation. Upon substitution in (1) we merge the forecast error into the equation error and include an intercept to yield the equation for estimation as:

$$
\pi_t = \gamma_0 + \lambda x_t + \gamma_f \pi_{t+1} + \gamma_b \pi_{t-1} + \epsilon_t, \tag{3}
$$

where the endogeneity problem arises from possible correlation of ϵ_t errors with the regressors. Hence it is well known that ordinary least squares (OLS) estimates $\hat{\gamma}_0, \hat{\lambda}, \hat{\gamma}_f, \hat{\gamma}_b$ of the parameters in (3) might be inconsistent. Galí and Gertler (1999) and many others assume that $\gamma_0 = 0$, that is force the line of regression through the origin.

Econometrics texts including Vinod (2008) explain super-consistency of OLS in the presence of nonstationary (integrated $I(1)$ of order 1) variables on two sides of the regression. However things become less neat when there are $I(d)$ regressors containing a range of $d \in [0, 1]$, say, possibly fractional values. Similarly, textbooks explain that OLS is inconsistent and might have to be replaced by the GMM if regressors are correlated with regression errors due to endogeneity. Unfortunately, actual sizes of such correlations and possible weakness of available GMM instruments remain unknown in given data. Hence, one generally does not know if OLS is seriously flawed and must be replaced by the GMM.

This paper studies local functioning of OLS and GMM estimators in the neighborhood of the available sample size n. Most economists would agree that our quarterly data over the past 70-year period is finite but sufficiently long. Although one can choose even longer data series to achieve better asymptotics, underlying macroeconomic notions might become too stretched. For example, the operational meaning of inflation variable (GDP deflator) to economic agents cannot he assumed to be constant, while the quality and content of the GDP change dramatically over time.

Assuming that OLS is inconsistent, Galí and Gertler (1999), KM09, among others, use GMM with potentially weak instrumental variables (up to four lags of inflation, the labor income share, the output gap, the long-short interest rate spread, wage inflation, and commodity price inflation) to overcome the endogeneity problem. This paper suggests an evidence-driven alternative.

Krugman (2009) argues that macro economists failed to predict the great recession of 2008 and various bubbles in asset prices because they have been "mistaking beauty for truth," where the beauty refers to mathematical elegance. Accordingly, this paper suggests a case-by-case computer intensive evaluation of consistency of individual regression coefficients, undeterred by the absence of beautiful asymptotic econometrics for models with mixtures of potential problems.

Section 2 reviews some notions from the maximum entropy (ME) bootstrap for time series and some tools in R for learning mathematical theory of convergence concepts. We use these tools to approximately evaluate the consistency of OLS estimates of (3). Section 3 contains our numerical example along with the R code for NKPC and Section 4 contains Efficient Estimation of NKPC Model. Section 5 considers estimation and confidence intervals for the three deep parameters and Section 6 contains our final remarks.

2 R packages called 'meboot' and 'ConvergenceConcpets'

Efron's bootstrap for independent and identically distributed (iid) data is an established computer intensive tool. In the context of time series, it applies readily to stationary data. Singh (1981) pointed out the inadequacy of iid bootstrap for stationary but dependent time series (m-dependent) data. Liu (1988) and Lahiri (2003) study bootstrap for non-iid but stationary situations in detail. Vinod and López-de-Lacalle (2009) provide an R package called 'meboot' for implementing the maximum entropy bootstrap from Vinod (2004) for strongly time-dependent non-stationary data.

Our NKPC of (3) potentially involves a mixture of nonstationary I(d) series, where the order of integration d can be uncertain and perhaps even fractional (long memory). Over the 70-year time period there may have been inflation regime switching structural changes. In addition to avoiding stationarity, 'meboot' package is designed to avoid the following three properties of traditional iid bootstrap mentioned in Vinod (2004, 2006). In this section it is convenient to let x_t denote a generic variable for which bootstrap resamples are needed.

- The traditional bootstrap sample obtained from shuffling with replacement repeats some x_t values while not using as many others. It never admits nearby data values in a resample. We are considering applications where there is no reason to believe that values near the observed x_t are impossible. For example, 1967 GDP deflator during the first quarter was 23.612, and we know that 23.6124 and 23.6119 round to 23.612. There is no justification for excluding all such nearby values.

- The traditional bootstrap resamples must lie in the closed interval $[\min(x_t), \max(x_t)]$. Since the observed range is random, we cannot rule out somewhat smaller or larger x_t.

- The traditional bootstrap resample shuffles x_t such that any depen-

dence information in the time series sequence $(x_1, \ldots, x_t, x_{t+1}, \ldots, x_T)$ is lost in the shuffle. If we try to restore the original order to the shuffled resample of the traditional bootstrap, we end up with essentially the original set x_t, except that some dropped x_t values are replaced by the repeats of adjacent values. Hence, it is impossible to generate a large number J of sensibly distinct resamples with the traditional bootstrap shuffle without admitting nearby values, which is what 'meboot' does.

Vinod and López-de-Lacalle (2009) explain the 'meboot' computer intensive algorithm with examples and graphics. Given a time series x_t, where $t = 1, 2, \ldots T$, the algorithm creates $j = 1, 2, \ldots J$ random resamples: $x_{t,j}$. Briefly, the steps are as follows.

1. Construct a matrix A with $t = 1, \ldots T$ in the first column and x_t in the second column. Now sort both columns simultaneously on the second column to yield order statistics $x_{(t)}$ in the second column.

2. Compute a neighborhood of each $x_{(t)}$ as the interval $[z_{t-1}, z_t]$, where $z_t = (x_{(t)} + x_{(t+1)})/2$ for $t = 1, \ldots, T-1$ are called intermediate points. We omit description of z_0 and z_T defining the neighborhoods of the first and last order statistic for brevity. The 'mass preserving constraint' ensures that each neighborhood of the data point x_t defined from these intermediate points has probability $1/T$ of being selected in the sample. This is similar to the iid boot.

3. Compute the mean of the maximum entropy density within each interval such that the 'mean-preserving constraint' (designed to eventually satisfy the ergodic theorem) is satisfied.

4. Generate random numbers from the $[0, 1]$ uniform interval, compute sample quantiles of the ME density at those points, sort them and place them in the second column of the matrix A.

5. This time, sort on the first column of A to create the sequence $1, 2, \ldots T$ while simultaneously sorting on the second column. The resulting values in the second column represent our 'meboot' j-th resample.

6. Repeat steps 3 to 5 J times (The algorithm default is $J = 999$).

Vinod (2010) views the double sorting algorithm as a mapping of x_t series to $x_{(t)}$ from the time domain to numerical magnitudes or values domain (v-dom). When only the range of the random variable is known, and we wish to be maximally non-committal about its functional form, maximum entropy (ME) principle suggests the uniform density. In our context the ME density is a patchwork of uniform densities residing in the v-dom defined over small neighborhoods defined in Step 2 above. Joag-dev (1984) surveys the literature on stochastic dependence and notes that dependence between two random measurements X and Y is high when their order statistics match. The algorithm matches the order statistics of each resample $x_{t,j}$ with the order statistics $x_{(t)}$ of the original data. Various examples in Vinod and López-de-Lacalle (2009) show that the algorithm is easy to use in practice.

The 'meboot' package allows us to create a large number (=J) of reincarnations of π_t, x_t variables which are indeed similar, but not too similar. Their autocorrelation and partial autocorrelation functions (acf, pacf) are also similar without using common parameters. See Vinod and López-de-Lacalle (2009) for an illustration of plots of (original and resampled) airline passenger series with acf, pacf and power spectra. The plots show that 'meboot' resamples are plausible representations of what might happen to the series, despite the presence of shifting seasonalities and regime changes.

Statistical inference for time series assumes that the observed series is one realization ω from the infinite dimensional ensemble $\omega \in \Omega$ of all possible series. The 'meboot' package provides a computer intensive approximation to Ω. Khintchine's "law of iterated logarithm" provides that assuming bounded means and variance, iid distributions, paths of partial sum sequences can fluctuate within a bound $\pm\sqrt{[2\sigma^2 n log\,(log(n))]}$, where n is sample size and σ^2 is the variance.

Now we review the R package called 'ConvergenceConcepts,' Lafaye de Micheaux and Liquet (2009) (hereafter, CC) inteneded for learning asymptotic theory by allowing students to work with concrete examples. The CC package directly studies similar paths of partial sum sequences $(X_{n,\omega} - X_\omega)$ based on n observations as $n \to \infty$. We write convergence in probability for sample size n as: $X_n \xrightarrow{p} X$ whenever

$$\forall \epsilon > 0 \;\; p_n = P[\omega;\, |X_{n,\omega} - X_\omega| > \epsilon] \to 0 \text{ as } n \to \infty. \tag{4}$$

The CC package explicitly evaluates p_n as a relative frequency. It counts the sample paths that wander outside the $\pm\epsilon$ bound among a large number $J(=999)$ of sample paths indexed by ω.

Almost sure convergence means that the limit of $X_{n,\omega}$ equals X_ω for all paths as $n \to \infty$. The CC package also reports a similar relative frequency of the count of the number of times that a $k \geq n$ exists such that $|X_{n,\omega} - X_\omega| > \epsilon$. The relative frequency provides an approximation to the probability a_n for 'almost sure' convergence. More important, the CC package permits us to approximately evaluate whether the relative frequencies satisfying the inequality $p_n < a_n$ are individually decreasing as n increases, in the context of a particular problem.

Consider a bootstrap approximation to a pivotal regression statistic, assumed to be a smooth functional of the true unknown cumulative density, conditioned on the available sample. Hall (1992) describes mathematical details of bootstrap convergence concepts. Using a generic notation, let our pivotal statistic be b_i denoting the ordinary least squares (OLS) estimate of the i-th true unknown coefficient β_i. The bootstrap resampling here "conditions" on the observed sample to create J regression problems with bootstrap estimates denoted by b_i^*. The bootstrap is successful because the empirical distribution function of $b_i^* - b_i$ is often a good approximation to that of $b_i - \beta_i$. In a simulation study, β_i are known and we can directly study whether the bootstrap b_i^* converges to β_i as $n \to \infty$, provided we can vary n over a range of values, eliminating the middleman b_i, as it were. Thus b_i^* based on n observations play the role of $X_{n,\omega}$ and β_i play the role of X_ω in the notation of

the CC package.

If one chooses $\epsilon > 0$ too large, then too many paths are likely to lie within the $\pm\epsilon$ bound, making the probability of a path going outside the bounds approach zero. Conversely, if $\epsilon > 0$ is too small, too many paths readily go outside the bound making the probability of a path going outside the bounds approach unity. Since we need to avoid both situations, a choice $\epsilon = 0.01$ for all coefficients does not work. Fortunately, the simulation allows us to choose our ϵ as $\bar{\epsilon} = \overline{(|b_i^* - \beta_i|)}$, where the average is over a large number of simulated sample paths admitting various sample sizes. These $\bar{\epsilon}$ values are expected to be different for each regression coefficient.

The following section illustrates our use of these packages for the NKPC example.

3 Numerical NKPC example

Now we check whether the coefficient estimates $\hat{\gamma}_0, \hat{\lambda}, \hat{\gamma}_f, \hat{\gamma}_b$ of (3) converge 'in probability' and /or 'almost surely.' We run the OLS regressions with and without the intercept. Since the intercept is statistically insignificant it might be appropriate to force the line of regression through the origin. However, the estimate of $\hat{\lambda}$ remains insignificant in both cases. Since the NKPC specification requires us to retain λ in the model, we have the option to retain the intercept.

Table 1 reports our OLS results. The adjusted coefficient of determination is: $R^2 = (0.6728, 0.8819)$ for the models with and without the intercept, respectively. The p-values for the F tests for significance of overall models are close to zero for both models.

Now we report the results based on the GMM estimated with and without the intercept, and two choices of of instrumental variables. Notation 'gmm2' refers to the instrument choice π_{t-3}, π_{t-4}. Notation 'gmm3' refers to the choice x_{t-3}, x_{t-4}. The GMM estimates from the R package 'gmm,' Chausse (2009), yield exactly the same estimates as the OLS when the instrumen-

Table 1: OLS regression results with intercept (upper panel) and without intercept (lower panel) for the NKPC model of equation (3)

Parameter	Estimate	Std. Error	t value	Pr(>|t|)
γ_0	7.4868	4.1674	1.80	0.0737
λ	6.4049	9.3445	0.69	0.4938
γ_f	0.4580	0.0497	9.21	0.0000
γ_b	0.4453	0.0498	8.94	0.0000
λ	8.9282	9.2824	0.96	0.3371
γ_f	0.4857	0.0475	10.22	0.0000
γ_b	0.4732	0.0475	9.95	0.0000

tal variables are absent, as they should. We use the 'gmm' package with (wmatrix="optimal") for the weight matrix appearing in the GMM objective function and with the 'continuous updated GMM'

Table 2: GMM results with and without intercept for the NKPC model of equation (3) where gmm2 has π_{t-3}, π_{t-4} as instruments and where gmm3 has x_{t-3}, x_{t-4} as instruments

estimator	$\hat{\gamma}_0$	$\hat{\lambda}$	$\hat{\gamma}_f$	$\hat{\gamma}_b$
gmm2.with.intercept	15.2830	15.5568	0.4879	0.3632
gmm2.NO.intercept		8.1709	0.4739	0.4921
gmm3.with.intercept	8.1363	6.5145	0.4564	0.4423
gmm3.NO.intercept		8.8153	0.4844	0.4747

KM09 impose the restriction that $\gamma_f + \gamma_b = 1$. It appears that in Table 2 where no restriction is imposed, the sum of these coefficients is always a bit less than unity.

The main motivation for using the GMM is to remove possible inconsistency of OLS induced by the endogeneity of regressors. Now we use a

slightly improved version of 'checkConv' function in 'meboot' package to approximately check the convergence of OLS in our current situation. A modification is needed, since leads and lags of the dependent variable π_t are present on the right hand side of (3). The simulation needs to correctly use these leads and lags.

There are $N = 236$ observations in the data set after omitting suitable lagged values. We reserve the last 100 data points for the study of convergence properties and let n reside range in the closed interval [137, 236] for our 100 sample paths. Clearly, there is a separate path for each regression coefficient with its own $\bar{\epsilon}$ available in our simulation environment based on average errors over these 100 estimates.

We use the CC package to evaluate the probabilities p_n defined in equation (4) for 'convergence in probability' and a similar probability a_n for 'almost sure convergence.' We have 100 evaluations of p_n, a_n as n increases in the closed interval [137, 236]. Convergence means these probabilities should decline as n increases, or $(dp_n/dn) < 0$ and $(da_n/dn) < 0$. We can approximately evaluate these derivatives by running two regressions of p_n and a_n on n. If we have consistency, these regressions should have statistically significantly negative slopes, indicated by the p-vales of these slopes being 'small' (< 0.05) at the 5% level.

Table 3 reports an application of this method to our model (1) evaluating convergence of coefficients. This Table uses the simple linear regression of probabilities p_n and a_n on n, omitting the details such as values of the estimated intercept in these regressions.

Table 3 reveals whether t-values in a regression of relative frequencies of sample paths p_n, a_n on sample sizes n of those paths are large enough to indicate statistically significantly declining probabilities for the OLS estimator. All slopes are significantly negative, except for $\hat{\gamma}_b$, which remains significantly negative at the 10 percent level. Although our very partial information about the functioning of the OLS estimator suggests slight doubt with 'convergence in probability,' the evidence supporting convergence 'al-

Table 3: Table for convergence of OLS coefficients of NKPC model regressing on n (without the n^2 term)

	(p/as).coef.	Estimate	Std. Error	t value	Pr(>|t|)
1	prob. $\hat{\gamma}_0$	-0.012947	0.000461	-28.080136	0
2	a.s. $\hat{\gamma}_0$	-0.01304	0.000452	-28.87482	0
3	prob. $\hat{\lambda}$	-0.014789	0.000756	-19.55699	0
4	a.s. $\hat{\lambda}$	-0.014867	0.000776	-19.157708	0
5	prob. $\hat{\gamma}_f$	-0.001556	0.000141	-11.033583	0
6	a.s. $\hat{\gamma}_f$	-0.004574	7.4e-05	-62.006419	0
7	prob. $\hat{\gamma}_b$	-0.000188	0.000111	-1.696236	0.093016
8	a.s. $\hat{\gamma}_b$	-0.003776	8.9e-05	-42.488981	0

most surely' at the same 5% level seems to overcome it. After all, the almost sure convergence implies convergence in probability.

Now we investigate what happens if try to refine these results by considering nonlinear regressions of p_n and a_n on $1, n, n^2$, where 1 represents the intercept. We let $n \in [137, 236]$, as before. The results for convergence in probability are given in Table 4 with greater detail. Upon fitting the quadratic we are interested in evaluating the derivative (dp_n/dn) at $n = N = 236$, the ending size of the data set, even though the evaluation of the derivative is somewhat sensitive to where we evaluate it.

It is desirable for convergence that this rate of change in probability be negative. The value of this derivative will be different for each of the four coefficients of the NKPC model. Convergence of the estimate of each coefficient requires this derivative to be negative. The table reports fuller details of the quadratic fit with all quadratic coefficient estimates, standard errors, Student's t values and p-values or $\Pr(> |t|)$. We include three lines to report results when p_n is regressed on 1 (for the intercept), n and n^2. The three lines must be present for each of the four coefficients of the NKPC model. It is convenient to report the derivative (dp_n/dn) evaluated at $n = N = 236$

along the middle line (see in Table 4 line seq. values 2, 5, 8 and 11). Note that values at lines 8 and 11 are not negative, raising questions about the convergence in probability for $\hat{\gamma}_f$ and $\hat{\gamma}_b$.

The results for almost sure convergence in Table 5 report three lines for each coefficients, with the middle line (see line seq. values 2, 5, 8 and 11) showing evaluations of (dp_n/dn) at N. Here all derivative evaluations are negative suggesting almost sure convergence of the OLS estimator. This is consistent with the result based on the regression of probabilities on n.

Table 4: Table for convergence in probability of OLS coefficients of NKPC model regressing p_n on 1 (intercept), n and n^2, and $(dp_n)/dn$ at N

seq.	ref. coef.	Estimate	Std. Error	t value	Pr(>|t|)
1	$p_n(\hat{\gamma}_0)$ on1	3.39526	0.64116	5.29546	0.00000
2	at N:-0.01170, on n	-0.01776	0.00684	-2.59685	0.01087
3	$p_n(\hat{\gamma}_0)$ on n^2	0.00001	0.00002	0.70512	0.48243
4	$p_n(\hat{\lambda})$ on 1	1.64501	1.03965	1.58227	0.11684
5	at N:-0.01954, on n	0.00351	0.01109	0.31655	0.75227
6	$p_n(\hat{\lambda})$ on n^2	-0.00005	0.00003	-1.65405	0.10135
7	$p_n(\hat{\gamma}_f)$ on 1	2.30427	0.11018	20.91331	0.00000
8	at N:0.00288, on n	-0.01862	0.00118	-15.84603	0.00000
9	$p_n(\hat{\gamma}_f)$ on n^2	0.00004	0.00000	14.55520	0.00000
10	$p_n(\hat{\gamma}_b)$ on 1	1.58740	0.10319	15.38351	0.00000
11	at N:0.00293, on n	-0.01219	0.00110	-11.07737	0.00000
12	$p_n(\hat{\gamma}_b)$ on n^2	0.00003	0.00000	10.93194	0.00000

Table 5 is similar to Table 4 detailing a quadratic fit for a_n on 1, n and n^2. The middle rows report the derivatives evalueated at N for OLS. It is interesting that all evaluations are negative, suggesting approximate 'almost sure' convergence for OLS in Table 5.

Now we evaluate the convergence for the GMM estimator. As with the OLS, we use the simulation to define the band so that the proportion of paths

Table 5: Table for almost sure convergence of OLS coefficients of NKPC model regressing a_n on 1 (intercept), n and n^2, and $(da_n)/dn$ at N

seq.	ref. coef.	Estimate	Std. Error	t value	Pr(>|t|)
1	$a_n(\hat{\gamma}_0)$ on 1	3.19808	0.62916	5.08306	0.00000
2	at N:-0.01243, on n	-0.01540	0.00671	-2.29432	0.02393
3	$a_n(\hat{\gamma}_0)$ on n^2	0.00001	0.00002	0.35193	0.72565
4	$a_n(\hat{\lambda})$ on 1	1.44589	1.06378	1.35920	0.17723
5	at N:-0.02024, on n	0.00579	0.01135	0.51070	0.61072
6	$a_n(\hat{\lambda})$ on n^2	-0.00005	0.00003	-1.82526	0.07104
7	$a_n(\hat{\gamma}_f)$ on 1	1.95663	0.09022	21.68648	0.00000
8	at N:-0.00323, on n	-0.00974	0.00096	-10.12497	0.00000
9	$a_n(\hat{\gamma}_f)$ on n^2	0.00001	0.00000	5.38453	0.00000
10	$a_n(\hat{\gamma}_b)$ on 1	0.50946	0.08448	6.03052	0.00000
11	at N:-0.00624, on n	0.00572	0.00090	6.34899	0.00000
12	$a_n(\hat{\gamma}_b)$ on n^2	-0.00002	0.00000	-10.56357	0.00000

going outside the band: $\pm\bar{\epsilon}$ is then numerically computed. A comparison of $\bar{\epsilon}$ values based on average errors in our simulation for the four coefficients is found in Table 6. Note that the simulation error in the estimation of the intercept are very large for GMM compared to OLS. By contrast the error for $\hat{\lambda}$ are large for OLS compared to GMM. The intercept seems to affect the magnitude of all errors. However, we use these suitably chosen distinct $\bar{\epsilon}$ values when we use the CC package for OLS or GMM. Of course, we want the result to be true for all ϵ values. Since this is not feasible in practice, we have to choose some $\bar{\epsilon}$ value such that the result would be true for all $|\epsilon| > |\bar{\epsilon}|$.

What about convergence of GMM? Again, we use lagged π and x as instruments. The 'true' coefficient values are determined by initial GMM estimates for the purpose of the simulation of sample paths.

Table 7 is similar to Table 3. A study of Table 7 reveals whether t-

Table 6: A comparison of $\bar{\epsilon}$ values based on average simulated errors

Type	$\hat{\gamma}_0$	$\hat{\lambda}$	$\hat{\gamma}_f$	$\hat{\gamma}_b$
OLS	2.87	11.52	0.005	0.004
GMM	231.76	1.17	0.25	0.56

Table 7: Table for convergence of GMM coefficients of NKPC model regressing both p_n and a_n on intercept and n

	(p/as).coef.No	Estimate	Std. Error	t value	Pr(>|t|)
1	prob. $\hat{\gamma}_0$	0.004905	0.003673	1.335426	0.274019
2	a.s. $\hat{\gamma}_0$	-0.001401	0.000504	-2.781518	0.068904
3	prob. $\hat{\lambda}$	-0.009309	0.004999	-1.862112	0.159496
4	a.s. $\hat{\lambda}$	-0.009109	0.003807	-2.392804	0.096497
5	prob. $\hat{\gamma}_f$	0.002603	0.008074	0.322358	0.76834
6	a.s. $\hat{\gamma}_f$	-0.013213	0.003861	-3.421941	0.041784
7	prob. $\hat{\gamma}_b$	-0.053053	0.011335	-4.68044	0.018428
8	a.s. $\hat{\gamma}_b$	-0.053053	0.011335	-4.68044	0.018428

values in a regression p_n or a_n on n based on relative frequencies of sample paths straying outside the bounds are large enough to indicate statistically significantly declining probabilities for the GMM estimator. As before, Table 7 omits the details regarding the estimated intercept and reports only on the slope coefficients for brevity. Since the p-values do exceed 0.05 along the first five rows of the table with two values positive, the convergence of GMM is more seriously in doubt compared to the convergence of OLS in our analogous Table 3.

Table 8 for the GMM reports regression of p_n on $1, n, n^2$ and derivatives da_n/dn at $n = N$ (see line seq. values 2, 5, 8 and 11). Note that three out of four positive derivatives (dp_n/dn) suggest possibly serious problem of divergence 'in probability,' for GMM compared to one in Table 4 for OLS.

Table 8: Table for convergence in probability of GMM coefficients of NKPC model regressing p_n on 1 (intercept), n and n^2, and $(dp_n)/dn$ at N

seq.	ref. coef.	Estimate	Std. Error	t value	Pr(>|t|)
1	$p_n(\hat{\gamma}_0)$ on 1	58.0595	203.9843	0.2846	0.8027
2	at N:0.0092, on n	-0.4970	1.7435	-0.2851	0.8024
3	$p_n(\hat{\gamma}_0)$ on n^2	0.0011	0.0037	0.2879	0.8005
4	$p_n(\hat{\lambda})$ on 1	382.3900	90.4348	4.2283	0.0516
5	at N:0.0184, on n	-3.2551	0.7730	-4.2112	0.0520
6	$p_n(\hat{\lambda})$ on n^2	0.0069	0.0017	4.1992	0.0523
7	$p_n(\hat{\gamma}_f)$ on 1	-556.0178	234.2183	-2.3739	0.1409
8	at N:-0.0380, on n	4.7542	2.0019	2.3748	0.1408
9	$p_n(\hat{\gamma}_f)$ on n^2	-0.0102	0.0043	-2.3735	0.1409
10	$p_n(\hat{\gamma}_b)$ on 1	788.0883	335.0808	2.3519	0.1430
11	at N:0.0036, on n	-6.6785	2.8640	-2.3319	0.1450
12	$p_n(\hat{\gamma}_b)$ on n^2	0.0142	0.0061	2.3134	0.1468

What about almost sure convergence of GMM? Table 9 has two positive derivative evaluations for GMM suggesting 'almost sure' divergence, compared to none in Table 5 for the OLS. The absence of almost sure consistency (divergence) of GMM compared to OLS is a robust result. We have used another set of regressions where we regress on $1, n, (1/n)$ and again evaluated da_n/dn at $n = N$ to find that all four coefficient evaluations for OLS are negative, whereas only two evaluations for GMM are negative. Detailed tables for the alternate nonlinear specification are omitted for brevity.

Hence the available partial information about the functioning of estimators using a quadratic function of n further confirms that that the GMM estimator provides no convergence in probability advantage over OLS in the specific context of our NKPC estimation.

When the available time series are close to random walk (nonstationary, integrated of order 1), many econometrics texts including Vinod (2008) (See

Table 9: Table for almost sure convergence of GMM coefficients of NKPC model regressing a_n on 1 (intercept), n and n^2, and $(da_n)/dn$ at N

seq.	ref. coef.	Estimate	Std. Error	t value	Pr(>|t|)
1	$a_n(\hat{\gamma}_0)$ on 1	-22.6667	23.2273	-0.9759	0.4321
2	at N:-0.0031, on n	0.1994	0.1985	1.0042	0.4210
3	$a_n(\hat{\gamma}_0)$ on n^2	-0.0004	0.0004	-1.0113	0.4183
4	$a_n(\hat{\lambda})$ on 1	288.3878	75.5903	3.8151	0.0623
5	at N:0.0118, on n	-2.4518	0.6461	-3.7949	0.0630
6	$a_n(\hat{\lambda})$ on n^2	0.0052	0.0014	3.7808	0.0634
7	$a_n(\hat{\gamma}_f)$ on 1	-231.2743	142.4961	-1.6230	0.2461
8	at N:-0.0304, on n	1.9945	1.2179	1.6376	0.2432
9	$a_n(\hat{\gamma}_f)$ on n^2	-0.0043	0.0026	-1.6485	0.2410
10	$a_n(\hat{\gamma}_b)$ on 1	788.0883	335.0808	2.3519	0.1430
11	at N:0.0036, on n	-6.6785	2.8640	-2.3319	0.1450
12	$a_n(\hat{\gamma}_b)$ on n^2	0.0142	0.0061	2.3134	0.1468

page 187) note that the OLS estimate converges to the true value at the rate n compared to the usual rate $\sqrt{n}$. This is the super-consistency mentioned earlier. Our tables showing that OLS has better convergence than GMM in the context of NKPC model may be due to the super-consistency property of OLS involving integrated variables (π_t, x_t) outweighing the endogeneity effects. Of course, we have a combination of proxy and instrument choices and differeing degrees of nonstationarity (d of $I(d)$) of variables appearing in the NKPC model. It is impossible to know whether super-consistency or endogeneity are more important, except by tailor-made evaluations of the actual functioning of the estimators using simulations such as ours.

4 Efficient Estimation of NKPC Model

While OLS may be converging, we know that the presence of autocorrelated regression errors can lead to inefficient estimates. Generalized least squares (GLS) estimation allows for non-spherical regression errors in eq. (3). Some R tools for this discussed in Vinod (2010) are applied in this section to the NKPC model. We use an R function called 'bestArma' to help decide which ARMA(p,q) model fits the regression residuals best.

We study Akaike Information Criterion (AIC) for various ARMA(p,q) models ranked by the AIC values from the smallest to the largest. For annual data, one does not usually choose the AR order larger than 2. Since this is quarterly data, it is reasonable to allow for order 4 error autocorrelations. The best model associated with the smallest AIC represents ARMA(4, 2). Its coefficients with standard errors are:

```
          ar1     ar2     ar3     ar4      ma1      ma2
      -0.6484  0.4989  0.6164  0.3980  -0.0349  -0.8096
s.e.   0.1021  0.0822  0.0724  0.0624   0.0958   0.0913
```

The GLS coefficient estimates for the NKPC model are: $\lambda = 0.1827763, \gamma_f = 0.5226623, \gamma_b = 0.5199869$. Correcting for the first order serial correlation is fairly common. Using GLS to correct for high order regression errors of the type ARMA(4, 2) explained in Vinod (2010) is somewhat new.

5 Inference for Deep Parameters

Now we turn to the problem of inference regarding the deep parameters (2) obtained by solving three highly nonlinear equations in three unknowns. KM09 are concerned with the identification of parameters in the NKPC model. We propose using Godambe (1985) pivot function (GPF) relying on his theory of estimating functions explained in Vinod (2008, Sec. 10.3).

The estimating functions are similar to 'moment conditions' familiar in

the context of GMM estimation. The GPF is defined as

$$\text{GPF} = \sum_{t=1}^{T} g_t^* \Bigg/ \left[\sum_{t=1}^{T} E(g_t^*)^2 \right]^{1/2}, \tag{5}$$

where g_t^* is the 'scaled quasi-score function' from the underlying quasi likelihood function also known as the optimal estimating equation.

The GPF avoids problematic Wald-type pivotal statistic commonly used in the usual t-tests altogether. It is the standard error in their denominators, which is known to contribute to possible lack of identification as shown in Dufour (1997) also cited by KM09. Vinod (1997) extends the GPF to multivariate regression problems. When there are T observations, in the simpler scalar case he rewrites the GPF as a sum of T scaled quasi-scores:

$$\text{GPF} = \sum_{t=1}^{T} S_t/S_c = \sum_{t=1}^{T} \tilde{S}_t, \quad \text{where } S_c = \left[\sum_{t=1}^{T} E(S_t)^2 \right]^{1/2}, \tag{6}$$

where we denote scaled quasi-score functions as: $\tilde{S}_t$. As a sum of T items, the central limit theorem assures us that GPF $\sim N(0,1)$, is asymptotically unit normal. Thus, the probability distribution of GPF is independent of unknown parameters and therefore it is a pivot.

Vinod (2008) (See pages 456-457) provides the code for the R function 'gpf' which computes GPF confidence intervals for each regression coefficient in a multiple regression sequentially. It involves an application of the Frisch-Waugh theorem. Our R code uses the 'gpf' and the R package called 'rootSolve' to solve the three nonlinear equations involved in deep parameters.

The estimate of deep parameters based on original OLS estimates (Table 1 upper panel with intercept and lower panel without intercept) without any ARMA adjustment are respectively: $\alpha = (0.0751, 0.0532)$ for the probability that prices remain fixed, $\beta = (0.8166, 0.9184)$ for the discount factor, and $\phi = (0.0596, 0.0476)$, the fraction of backward looking price setters in the economy.

Note that if we use efficient GLS estimates after adjusting for ARMA(4, 2) regression errors, the deep parameters become $\alpha = 0.47, \beta = 1.24, \phi = 0.58$. See the column entitled 'GLS' in Table 10. We also report Godambe Pivot Function (GPF) confidence intervals for the estimates of deep parameters restricted to the range [0,1] (not always binding).

Table 10: GPF confidence limits on deep parameters for OLS, OLS$_{no.int}$ (no intercept) and GLS after correcting for ARMA(4, 2) regression errors

Parameter	OLS	OLS$_{no.int}$	GLS	Lower	Upper
α	0.08	0.05	0.47	0.21	0.99
β	0.82	0.92	1.24	0.00	1.00
ϕ	0.06	0.05	0.58	0.29	0.72

Thus we have shown how to use the confidence limits on slopes in eq. (3) to construct confidence limits on deep parameters of economic interest. For example, the GLS estimate of ϕ, the fraction of firms looking backward when they set their prices at 58% with a confidence range of 29 to 72 percent appears to be intuitively plausible to me than the OLS value of 8%. Similarly the GLS point estimate $\hat{\alpha} = 0.47$ for the probability that prices remain fixed with the indicated range also seems more plausible than the OLS value of 0.06. Bils and Klenow (2004) study frequency of price changes for 350 categories of goods and services covering about 70 percent of consumer spending, using unpublished 1995-97 data from the Bureau of Labor Statistics (BLS). Even though they observe dramatic variation in the frequency of price changes across goods, they find that about half of prices remain unchanged for 5.5 months or less. Christiano et al. (2005) have a model explaining the observed inertia in inflation and output such as ours. These and other results in the literature suggest that our GLS estimates of α and ϕ are plausible.

Similar to KM09, our confidence interval estimation does impose the range limit [0, 1] on all deep parameters. The range limit binds only our estimate of the discount factor β, implying that our confidence interval for

the discount rate is problematic. The OLS point estimate is $\hat{\beta} = 0.82$, but the GLS point estimate exceeding unity is perhaps meaningless, even though slightly negative interest rate for time-preference calculations in the US is possible to imagine. Together, these estimates suggest that the point estimate of β is close to 1, even though we do not impose explicit constraint on it. Recall that KM09 impose the restriction $\beta = 1$, which in turn imposes the restriction that $\gamma_f + \gamma_b = 1$.

The continuous updated GMM estimates of 95% confidence intervals in Table 3 of KM09 are $\alpha = [0.56, 1]$ and $\phi = [0, 1]$ with no confidence interval for β. Their GMM interval is binding at the upper limit for α and both limits for ϕ. Clearly, our efficient GLS estimates of all deep parameters are more attractive than GMM estimates in KM09, perhaps because GMM estimates might not be converging.

6 Final Remarks

A need for greater realism in macro-econometric work is obvious in light of recent failure to predict the great recession of 2008. We need a fresh look at unique challenges arising from evolving dynamics, simultaneity, structural change of economic regimes and changing nonlinearities. The fresh look in this paper exploits the free R software. In modern medicine, there is an attempt to tailor-make the medicines to individual patient using genetic and other specific information. We use the CC package in R to tailor-make an inference for the specific non-stationary time series with all their individual quirks, similar to the 'warts and all' of an individual patient.

Our illustration focuses on the New Keynesian Phillips Curve model of equation (1) specified for estimation as equation (3). It is commonly estimated by various types of GMM estimators in the literature, with an array of possibly weak instrumental variables. We find that many GMM estimates using recent quarterly US data are not necessarily superior to the OLS or GLS.

The CC package in R allows us to evaluate sample path probabilities p_n for 'convergence in probability' and a_n for 'almost sure' convergence in the context of a simulation based on equation (3). This allows us to check whether p_n, a_n probabilities indeed decline as n increases over a range of 100 evaluations of 999 sample paths. A statistically significant slope coefficient of p_n, a_n regressed on n should be negative for sample path evidence of convergence.

The key reason for using the GMM is possible removal of inconsistency arising from endogeneity of regressors. A simulation reveals that a large number of sample paths of associated with some GMM coefficients might actually diverge. Upon considering further evidence based on quadratic fits, we find that GMM offers no convergence advantage over OLS in the context of our model. This may be because the GMM might be removing endogeneity, but OLS retains its super consistency in the presence of possibly $I(1)$ variables already present in the NKPC model.

Certain identification problems associted with deep parameters are known in the NKPC literature. We use Godambe's pivot function (GPF) from Biometrics to construct confidence intervals for our deep parameters of the NKPC model, because that pivot is a sum of N scaled scores when there are N observations, permitting a direct use of the central limit theory, while overcoming identification problems.

Our GPF confidence intervals have reasonable values for two of the three deep parameters. Hence our GPF intervals associated with efficient GLS estimation in our Table 10 seem to offer a superior alternative to Table 3 on page 306 of KM09, who needed to impose the constraint $\beta = 1$ and yet report that the upper bound of their 95% interval for $\alpha = [0.56, 1]$ is indeterminate. Moreover their confidence interval for $\phi = [0, 1]$ is fully indeterminate, because binding range constraints are present.

References

Mark Bils and Peter J. Klenow. Some evidence on the importance of sticky prices. *Journal of Political Economy*, 112(5):947–985, 2004.

Pierre Chausse. *gmm: Generalized Method of Moments and Generalized Empirical Likelihood*, 2009. URL `http://CRAN.R-project.org/package=gmm`. R package version 1.1-1.

Lawrence J. Christiano, Martin Eichenbaum, and Charles L. Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45, 2005.

J-M Dufour. Some impossibility theorems in econometrics with applications to structural and dynamic models. *Econometrica*, 65:1365–1387, 1997.

J. Galí and M Gertler. Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, 44:195–222, 1999.

V P Godambe. The foundations of finite sample estimation in stochastic processes. *Biometrika*, 72:419–428, 1985.

P. Hall. *The bootstrap and Edgeworth expansion*. Springer Verlag, New York, 1992.

Kumar Joag-dev. Measures of dependence. In P. K. Krishnaiah and P. K. Sen, editors, *Handbook of Statistics, Vol. 4*, pages 79–88. North Holland, New York, 1984.

F Kleibergen and S. Mavroeidis. Weak instrument robust tests in gmm and new keynesian phillips curve. *Journal of Business and Economic Statistics*, 27(3):293–311, 2009.

Paul Krugman. How did economists get it so wrong? *New York Times Magazine*, September 2, 2009.

P. Lafaye de Micheaux and B. Liquet. Understanding convergence concepts: a graphical simulation based approach. *The American Statistician*, 63(2):173–178, 2009. URL `http://www.biostatisticien.eu/ConvergenceConcepts/`. Software: R Package, ConvergenceConcepts, version 0.5.

S. N. Lahiri. *Resampling Methods for Dependent data*. Springer Verlag, New York, 2003.

R. Y. Liu. Bootstrap procedures under some non-iid models. *The Annals of Statistics*, 16(4):1696–1708, 1988.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. URL `http://www.R-project.org/`. ISBN 3-900051-07-0.

K Singh. On the asymptotic accuracy of efron's bootstrap. *The Annals of Statistics*, 9:1187–1195, 1981.

H D Vinod. Using godambe–durbin estimating functions in econometrics. In I Basawa, V P Godambe, and R Taylor, editors, *Selected Proceedings of the Symposium on Estimating Equations*, volume 32 of *IMS Lecture Notes — Monographs Series*, pages 215–237, 1997.

H D Vinod. Ranking mutual funds using unconventional utility theory and stochastic dominance. *Journal of Empirical Finance*, 11(3):353–377, 2004.

H D Vinod. Maximum entropy ensembles for time series inference in economics. *Journal of Asian Economics*, 17(6):955–978, 2006.

H. D. Vinod. Superior estimation and inference avoiding heteroscedasticity and flawed pivots: R-example of inflation unemployment trade-off. In H. D. Vinod, editor, *Advances in Social Science Research Using R*, pages 39–63. Springer, New York, 2010.

H. D. Vinod and Javier López-de-Lacalle. Maximum entropy bootstrap for time series: The meboot R package. *Journal of Statistical Software*, 29(5): 1–19, 2009. URL `http://www.jstatsoft.org/v29/i05/`.

Hrishikesh D. Vinod. *Hands-on Intermediate Econometrics Using R: Templates for Extending Dozens of Practical Examples*. World Scientific, Hackensack, NJ, 2008. URL `http://www.worldscibooks.com/economics/6895.html`. ISBN 10-981-281-885-5.