CAHIER 9335

DYNAMIC REGRESSION AND FILTERED DATA SERIES:
A LAPLACE APPROXIMATION TO THE EFFECTS
OF FILTERING IN SMALL SAMPLES

Eric Ghysels¹ and Offer Lieberman²

¹ Département de sciences économiques and Centre de recherche et développement en économique (C.R.D.E.), Université de Montréal.
² Department of Economics, University of Bristol.

November 1993

* The first author would like to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada, the Fonds pour la formation de chercheurs et l’aide à la recherche (FCAR) and would like to thank the Department of Econometrics at Monash University for their generous hospitality. We would also like to thank Peter Phillips for insightful comments he made on an earlier draft as well as seminar participants at Yale and the Montréal Econometrics Workshop, in particular, Vicky Zinde-Walsh our discussant, and Sophie Mahseredjian who provided very helpful research assistance.
Ce cahier a également été publié au Centre de recherche et développement en économique (C.R.D.E.) (publication no 3293).

Dépôt légal - 1993
Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

ISSN 0709-9231
Résumé

Nous utilisons la méthode de Laplace pour évaluer l'effet de filtrage de données sur la distribution de l'estimateur mco dans des échantillons de taille finie. Notre étude se concentre principalement sur le biais de l'estimateur mco dans le modèle dynamique autorégressif d'ordre un avec régresseurs exogènes. Les formules de l'approximation de l'effet de filtrage (linéaire) sont très simples et requièrent uniquement des opérations d'algèbre matricielle. Nous illustrons la méthode au moyen du filtre X-11 et étudions sa précision.

Mots clés : modèles autorégressifs, biais d'échantillon fini, filtre X-11, ajustement pour les effets de saison.

Abstract

It is common for an applied researcher to use filtered data, like seasonally adjusted series, for instance, to estimate the parameters of a dynamic regression model. In this paper, we study the effect of (linear) filters on the exact finite sample distribution of parameters of a dynamic regression model with a lagged dependent variable and a set of exogenous regressors. So far, only asymptotic results are available. Our main interest is to investigate the effect of filtering on the small sample bias and mean squared error. In general, these results entail a numerical integration of derivatives of the joint moment generating function of two quadratic forms in Normal variables. The computation of these integrals is quite involved. However, we take advantage of the Laplace approximation to the bias and mse which substantially reduces the computational burden, as it yields relatively simple analytic expressions. We obtain an analytic formula for approximating the effect of filtering on the finite sample bias and mse. We evaluate the adequacy of the approximations by comparison with Monte Carlo simulations, using the X-11 filter as specific example.

Key words : autoregressive models, small sample bias, X-11 filter, seasonal adjustment filter.
1. INTRODUCTION

The single equation dynamic regression model:

\[ y_t = \alpha y_{t-1} + x_t^\prime \beta + \epsilon_t \]  

has been of interest in many econometric research programs, from the distributed lag model to its more recent use in the context Euler equation estimation or the vast literature on testing for unit roots. Of course, the nature of the assumptions on the regressors and disturbances have changed along these different programs.

While (1.1) is often a starting point for econometric analysis, it is also common for an applied researcher to use filtered data like seasonally adjusted series, for instance. The fact that filtered data are being used is typically ignored in theoretical considerations. There are some notable and well-known exceptions though. Sims (1974) and Wallis (1974), in particular, studied seasonal adjustment and regression analysis with exogenous regressors. They explored the nature of the asymptotic bias and the statistical properties of seasonal adjustment in regression when adjusted and unadjusted data were used. Recent work by Ghysels and Perron (1993), Hansen and Sargent (1993) and Sims (1993) has refocused attention on the subject. The emphasis of the first paper is slightly different from the other two. Namely, Ghysels and Perron are concerned with the effect of seasonal adjustment filtering on tests for a unit root, interpreting equation (1.1) as possibly having a constant and trend as exogenous regressors \( x_t \) and a general error process satisfying a set of regularity conditions [see, e.g., Phillips (1987)]. Hansen and Sargent, and Sims, on the other hand, dealt with rational expectations models with either exogenous sources of seasonality, discussed mostly by Sims, or endogenous mechanism generating seasonals, as considered by Hansen and Sargent.

In the models described above, the parameters of direct interest are not \( \alpha \) and/or \( \beta \) but "deeper" parameters mapping into \( \alpha \) and \( \beta \). Sims argued that using seasonally adjusted data, despite the potential severe biases and loss of information, may produce better estimates of the so-called deeper parameters. Hansen and Sargent endorsed such findings with a set of examples and developed a general asymptotic apparatus to deal with the question of interest. So far, most results for rational expectations models are based on a set of examples. For the case of unit root tests, we are perhaps on firmer ground, as there appears to emerge a general conclusion, namely that seasonal
adjustment procedures which produce biases in stationary models seem to have a power-reducing effect (in finite samples as well as asymptotically), since the biases are mostly upwards, at least for the standard and commonly used Census X-11 filter.

There is, of course, an extensive literature on the finite sample and asymptotic properties of the OLS estimator for $\alpha$ in equation (1.1). Particularly for the AR(1) model with or without constant and/or time trend, especially with $\alpha$ in the vicinity of or exactly equal to unity. Developments of the asymptotic distribution theory include Mann and Wald (1943), White (1958), Phillips (1977, 1978, 1987), Nabeya and Tanaka (1990), Perron (1991) and many others. Drawing on these results, Ghysels and Perron (1993) provide analytic characterizations of the effect of filtering on the asymptotic distribution of the OLS estimator of $\alpha$ both in the unit root and stationary cases. In this paper, a first and modest attempt is made to develop a general analytic characterization of filtering on the finite sample distribution of the OLS estimator for $\alpha$. While the developments of a small sample distribution theory are not as abundant as large sample results, there are nonetheless considerable contributions to draw from. All theoretical developments in finite samples revolve around the fact that, with Normal disturbances, the OLS estimator for $\alpha$ can be expressed as a ratio of quadratic forms in vectors of random Normal i.i.d. series. The major challenge is to find relatively simple expressions that make the effect of filtering on the finite sample behavior transparent enough to be expressed as analytic formulae.

For the general case, the exact moments of a ratio of quadratic forms in Normal variables are complicated functions of infinite sums of invariant polynomials with multiple matrix arguments, see Smith (1989). The strategy followed in this paper is to settle for approximations to the finite sample properties of the OLS estimator of $\alpha$. The method of approximation is one which needs to strike a delicate balance between (1) accuracy of approximation and (2) analytical simplicity of the approximation. An approximation which we propose as striking this balance, providing serviceable solutions to the study of filtering affects, is that of the Laplace approximation. The use of the Laplace approximation has been advanced by Phillips (1983), where it was proposed as a method of extracting marginal density approximations of instrumental variable estimators in a small sample theory context for the general single equation case. This method has also recently been adopted in the Bayesian statistics literature by, among others, Daniels and Young (1991), DiCiccio and Martin (1991), Tierney and Kadane (1986) and Tierney, Kass and Kadane (1989). Lieberman (1992) proposes and develops the Laplace approximation to the moments of a ratio of quadratic forms in
Normal variables and discusses the accuracy of the approximation. As the analytic expressions are relatively simple, they allow us to characterize the effect of filtering on the distributional properties of the OLS estimator explicitly, as desired. Indeed, the formulae we present characterizing the effect of filtering only involve linear algebra operations and hence do not require numerical integration as is often the case; see, for instance, Hoque and Peters (1986) and Magnus (1986).\footnote{An alternative approach not requiring numerical integration either was proposed by Kiviet and Phillips (1993). Except in some very special circumstances, their approach is not directly applicable here as it assumes uncorrelated errors.} Also, numerical integration to compute exact moments can often be cumbersome and ill-behaved. None of these problems appear in the approach pursued in this paper. We specifically focus on the X-11 filter case or, more precisely, the linear approximation to this standard seasonal adjustment procedure. We evaluate the adequacy of the Laplace approximation with Monte Carlo simulations, using the X-11 filter as specific example.

In section 2, notation, definitions and regularity conditions are introduced. Section 3 covers the Laplace approximation and its resulting formulae. The latter are used in section 4 to study the effect of the X-11 filter on estimation of \( \alpha \) in small samples. Section 5 concludes.

2. NOTATION, FILTERS AND ESTIMATORS

The purpose of this section is to introduce the notation and structure necessary to characterize and define the estimators. Consider again the dynamic regression model:

\[
y_t = \alpha y_{t-1} + x_t' \beta + \epsilon_t, \quad t = -L, \ldots, 0, 1, \ldots, T + k.
\]

Note that the sample starts at \( t = -L \) and ends at \( T + k \) to accommodate two-sided filtering of data. It is assumed that the regressors are nonstochastic and that the error process \( \epsilon' = (\epsilon_{-L}, \ldots, \epsilon_{T+k}) \) has a known covariance matrix, i.e., \( \text{E} \epsilon \epsilon' = \Omega \). The covariance matrix does not have to be diagonal, hence, we can allow for MA error process, for instance.

Suppose now that the data is filtered with a known linear filter. The most prominent example is that of the X-11 filter and its linear approximation, discussed by Young (1968), Wallis (1974) and Ghysels and Perron (1993), for instance. The filter weights of the possibly two-sided filter are represented by the vector \( f' = (f_{-L}, \ldots, f_k) \). On the basis of this vector, we define the \((T+1) \times (T+k+L+1)\) filter matrix \( F \) as
This matrix transforms a sample of size $T + k + l + 1$ into a filtered data set with $T + 1$ observations, since at each end of the sample, data are discarded by the two-sided filter. On the basis of (2.1) and (2.2), we further define filtered and unfiltered data sets as $y' = (y_{-l}, \ldots, y_{T+k})$, $y_F = F y$, $(T + 1) \times 1$ and $y_U = U y$, $(T + 1) \times 1$, where the matrix $U$ cuts away observations at each end of the sample, so that $y_F$ and $y_U$ are filtered and unfiltered data sets of equal sizes $T + 1$ drawn from the same random vector $\varepsilon$. Formally stated,

$$U \equiv (O_{(T+1)\times l} \quad I_{T+1} \quad O_{(T+1)\times k}).$$

Each element of $y_F$ and $y_U$ will be denoted $y^F_t$ and $y^U_t$, respectively, $t = 0, \ldots, T$. We write $\varepsilon_F = F \varepsilon$, $\varepsilon_U = U \varepsilon$, $X_U = UX$ and $X_F = FX$, where $X' = (x'_{-l}, \ldots, x'_{T+k}).$

While this setup is relatively simple, it should be noted that it is not unrealistic. In particular, one can think of a situation where a statistical agency releases data simultaneously, unadjusted and adjusted, covering the same sample period, while the agency has actually more raw data available than is released to the public. The setup also has another advantage which needs to be emphasized at this stage. The filter matrix $F$ appearing in (2.2) is not a square matrix. Consequently, filtering will always entail information loss that cannot be recovered or undone by an estimation procedure such as GLS which would essentially amount to taking the inverse of the matrix $F$.

The fact that $\Omega$ and $F$ are known matrices implies that the covariance matrices of the filtered and unfiltered errors are also known, namely:

---

2 In finite samples taken from a stationary data generating process, we know that the distribution of estimators such as the OLS estimator of $\alpha$ in (2.1) depends on the initial starting value $y_0$ (for references, see Introduction). The matrices $F$ and $U$, while creating an equal sample size, do not guarantee that the first elements of $y^F_t$ and $y^U_t$ are the same. In fact, the first elements of the two data vectors will almost surely be different.
(2.3a) $E_{\Omega_U} \Omega_U' = \Omega_U = U\Omega U'$

and

(2.3b) $E_{F} \Omega_F' = F\Omega F'$.

Upon setting

$$R_\alpha = \begin{bmatrix}
  b & 0 & 0 & \ldots & 0 & 0 \\
  b\alpha & 1 & 0 & \ldots & 0 & 0 \\
  b\alpha^2 & \alpha & 1 & \ldots & 0 & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
  b\alpha^{T-1} & \alpha^{T-2} & \alpha^{T-3} & \ldots & \alpha & 1
\end{bmatrix}, \quad b = \begin{cases}
  (1 - \alpha^2)^{-1/2} & \text{if } \alpha \in (-1, 1) \\
  0 & \text{otherwise}
\end{cases},$$

we have

(2.5a) $E y y' = \Omega_Y = R_\alpha \Omega R_\alpha', \quad \mathcal{T} \times \mathcal{T}$

(2.5b) $E y_F y'_F = \Omega_Y F R_\alpha \Omega R_\alpha', \quad (T + 1) \times (T + 1)$

and

(2.5c) $E y_U y'_U = \Omega_Y U R_\alpha \Omega R_\alpha', \quad (T + 1) \times (T + 1)$.

The matrix $R_\alpha$ is formulated such that it is assumed that $y_{-j}$ is drawn from the unconditional distribution of the $\{y_i\}$ process.\(^3\)

Let $D_1 = [1_{T}, \Omega_{T1}], D_2 = [\Omega_{T1}, 1_{T}]$ and $Z_j = D_j X_j, j = F$ or $U$. Then the OLS estimator of $\alpha$ and $\beta$ may be written as:

(2.6a) $\hat{\alpha}_j \equiv y'_j S_j y_j / y'_j B_j y_j$

and

(2.6b) $\hat{\beta}_j \equiv (Z_j' Z_j)^{-1} Z_j' (D_j y_j - \hat{\alpha}_j D_1 y_j)$

---

\(^3\) One exception, as noted in (2.4), is when $\alpha = 1$, for instance. In such a case, the starting value of the process does not affect the small sample distribution of the OLS estimator.
where

\[ S_j \equiv \frac{1}{2} (D_1^j A_j D_2 + D_2^j A_j D_1) \]

\[ B_j \equiv D_1^j A_j D_1 \]

\[ A_j = I - X_j (X_j^t X_j)^{-1} X_j^t. \]

The OLS estimator \( \hat{\alpha}_U \) has the obvious disadvantage that, apart from its small sample bias, it may not be consistent in large samples even if we start out with a white noise innovation structure for the DGP of the raw data process. In other words, even though \( \Omega \) may be diagonal, it is clear that \( \Omega_F \) will typically be nondiagonal. Consequently, it is not clear when we compare \( \hat{\alpha}_U \) and \( \hat{\alpha}_F \) whether we are studying the effect of filtering or the combination of lagged dependent variables in a regression with autocorrelated residuals. Since it is assumed that \( \Omega_F \) and \( \Omega_U \) are known, with \( \Omega \), \( U \) and \( F \) known, it is relatively straightforward to extend our study to GLS estimators.\(^4\)

Finally, we need to discuss some regularity conditions regarding the admissible set of filters as expressed by the filter weights \( f, t, \ldots, f, k \) in (2.2). The restrictions we impose are the same as in Ghysels and Perron (1993). It is important to note, however, that these restrictions, while crucial for the asymptotic theory to hold in the unit root case \( \alpha = 1 \), are not essential for the development of small sample results. Yet, the regularity conditions greatly simplify our analysis. Moreover, most filters which come to mind when one considers seasonal adjustment procedures feature the restrictions we are about to discuss. In particular, as Ghysels and Perron noted, the linear approximation to the X-11 procedure is among the admissible set of filters. The restrictions can be formulated as follows:

\[(2.7) \quad (a) \ t = k, \ (b) \ f_j = f_j \text{ for } j = 1, \ldots, k \text{ and } (c) \sum_{j=-t}^{k} f_j = 1.\]

With these restrictions holding, we can greatly simplify our results as often \( X_U \equiv X_F \) and consequently \( A_F \equiv A_U \) as well as \( B_F \equiv B_U \). Indeed, when the exogenous regressors

\(^4\) The extension to the GLS estimator is straightforward and will not be dealt with explicitly. In general, this estimator will not unravel the effect of filtering, because the F matrix is noninvertible.
consist of a constant and/or a trend, then the conditions in (2.7) result in this simplification. Again, it should be stressed that such restrictions are not necessary, yet they greatly simplify, for instance, the dependence of $\alpha_F$ on the filter matrix $F$ as $A$ and $B$ in (2.6a) are invariant to filtering.

3. THE LAPLACE APPROXIMATION AND THE EFFECT OF FILTERING ON BIAS AND MEAN SQUARED ERROR

The OLS estimators $\hat{\alpha}_j$, $j = F$ or $U$, introduced in the preceding section, are ratios of quadratic forms in Normal variables. For the general case, the exact moments of such a ratio are complicated functions of infinite sums of invariant polynomials with multiple matrix arguments, as discussed in Smith (1989). Alternative formulas which involve unsolved integrals are presented by Magnus (1986) and Sawa (1972). For our purposes, the exact results presented so far within the literature are analytically too complicated for the analysis of filtering and estimation. The bias and mean squared error of the estimators, filtered and unfiltered, entail a numerical integration of derivatives of the joint moment generating function of the two quadratic forms in hand and the filtering effect is not easy to isolate analytically, as will be made clear in this section.

Another route often taken to unveil the statistical properties of $\hat{\alpha}_j$ is to compute numerically the cumulative density function of the ratio of quadratic forms. Several algorithms have been suggested, including Imhof (1961), Davies (1973) or Shively et al. (1990). It is possible, for instance, to define a median unbiased estimator for $\hat{\alpha}_j$, as introduced by Andrews (1993) and to analyze the effect of filtering on it, but once again, the computational burden and lack of analytic expressions make this route unattractive as well.

As discussed in the Introduction, an alternative approach consists of utilizing approximations to the moments of this ratio via the Laplace method. As will be shown in this section, the approximations have the appeal of providing relatively simple analytic expressions, only involving matrix operations which are computationally straightforward with almost any standard econometrics software package. Moreover, it is relatively easy to isolate the effect of filtering with the Laplace approximation to the moments of the estimators $\hat{\alpha}_j^k$ with the resulting expressions.
We first briefly review the Laplace approximation, as derived by Lieberman (1992). Then, we discuss bias and mean squared error with and without filtering. More specifically, in Section 3.1, we briefly discuss the Laplace approximation to the ratio of quadratic forms. In Section 3.2, we derive the bias and mean squared error with filtered and raw series.

3.1 The Laplace approximation

For simplicity of the exposition, we first omit all subscripts on \( \hat{\alpha} \), \( y \), \( S \) and \( B \), returning to the previous notation once the general approximation is established.

The \( r^{th} \) moment of \( \hat{\alpha} \) involves inversion of the joint moment generating function of \( y'Sy \) and \( y'By \), being

\[
M(\omega_1, \omega_2) = E \exp(\omega_1 y'Sy + \omega_2 y'By).
\]

As in Phillips (1982, p. 35), the expression for the moments of \( \hat{\alpha} \) can be readily deduced to be

\[
E(\hat{\alpha})^r = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \frac{\partial^r M(\omega_1, \sum_{i=1}^{r} \omega_{2i})}{\partial \omega_1^r} \bigg|_{\omega_1=0} \, d\omega_{21} \ldots d\omega_{2r},
\]

where \( \omega_2 = \sum_{i=1}^{r} \omega_{2i} \). The first and second moments were given explicitly by Hoque and Peters (1986), but were retained as unsolved integrals.

To make (3.2) suitable for the application of Laplace's method, the last integral can be rewritten as

\[
E(\hat{\alpha})^r = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g_2(0, \sum_{i=1}^{r} \omega_{2i}) \exp(h(0, \sum_{i=1}^{r} \omega_{2i})) \, d\omega_{21} \ldots d\omega_{2r},
\]

where
\[
\begin{aligned}
g_r(0, \sum_{i=1}^r \omega_{2i}) &= \left\{ \frac{\partial \ln M(0, \sum_{i=1}^r \omega_{2i})}{\partial \omega_1} \bigg|_{\omega_1=0} \right\},
\end{aligned}
\]
and
\[
\begin{aligned}
h(0, \sum_{i=1}^r \omega_{2i}) &= \log M(0, \sum_{i=1}^r \omega_{2i}).
\end{aligned}
\]

The essence of the Laplace method is that when a key parameter tends to infinity, most of the contribution to the type of integrals in (3.3) arises from the immediate neighborhood of one point. Two distinct approximations are plausible, depending on whether the maximum of \(h(\cdot)\) is attained at an inferior point on the domain of integration, or at a boundary point. Lieberman (1992, Lemma 1) shows that when \(T \to \infty\), the integrand's value boundary points \(\omega_{2i} = 0, i = 1, \ldots, r\), and their immediate vicinity determine indeed most of the integral. Using this feature, it is shown that the Laplace approximation to the \(r^{th}\) moment of the estimator \(\hat{\alpha}_j\) is
\[
E_L(\hat{\alpha}_j)^r = \frac{E(y_j^* S_j y_j)^r}{(E y_j^* B_j y_j)^r},
\]
[regardless of the order of the moment, see Lieberman (1992, Theorem 1)]. The subscript \(L\) indicates that we are dealing with the Laplace approximation. With the full notation of the present paper, allowing for filtered and unfiltered data sets, the Laplace approximation for the first two moments of \(\hat{\alpha}_j, j = F, U\) are:

\[
(3.4) \quad E_L(\hat{\alpha}_j) = \frac{\mu_j^* S_j \mu_j + \text{tr}(\Omega_j^* S_j)}{\mu_j^* B_j \mu_j + \text{tr}(\Omega_j^* B_j)},
\]

\[
(3.5) \quad E_L(\hat{\alpha}_j)^2 = \frac{\left(\frac{\mu_j^* S_j \mu_j + \text{tr}(\Omega_j^* S_j)}{\mu_j^* B_j \mu_j + \text{tr}(\Omega_j^* B_j)}\right)^2}{\left(\frac{\mu_j^* S_j \mu_j + \text{tr}(\Omega_j^* S_j)}{\mu_j^* B_j \mu_j + \text{tr}(\Omega_j^* B_j)}\right)^2},
\]

where \(\mu_j = E(j y)_j = j R_{\alpha}^* X^* B\), with \(R_{\alpha}^*\) the same matrix as \(R_{\alpha}\) except that \(b\) is replaced in the first column by \((1 - \alpha)^{-1}\). The expressions (3.4) and (3.5) clearly reveal the appeal of using the Laplace approximation. The exact first and second moments involve the (numerical) integration of some complicated matrix formula. In contrast, equations (3.4) and (3.5) involve fairly simple and straightforward matrix operations such as calculating the trace of a matrix product.
3.2 Bias and mean squared error with filtered and raw data series – the Laplace approximation

We turn our attention now to characterizing the effect of filtering on the finite sample properties of our estimators via the Laplace approximations. First, we will present the general formula, followed by several specific cases that are of interest. From equations (3.4) and (3.5), the Laplace approximations to the bias and mean squared error of \( \hat{\alpha}_j \) are

\[
\text{Bias}_L(\hat{\alpha}_j) = \frac{\mu_j' (S_j - \alpha B_j) \mu_j + \text{tr}(\Omega^2_j (S_j - \alpha B_j))}{\mu_j' B_j \mu_j + \text{tr}(\Omega^2_j B_j)}
\]

\[
\text{MSE}_L(\hat{\alpha}_j) = \frac{\{\mu_j' S_j \mu_j + \text{tr}(\Omega^2_j S_j)\}^2 + 2\{\text{tr}(\Omega^2_j S_j)^2 + 2 \mu_j S_j \Omega_j S_j \mu_j\}}{\{\mu_j' B_j \mu_j + \text{tr}(\Omega^2_j B_j)\}^2 + \alpha \left[ \frac{\mu_j' S_j \mu_j + \text{tr}(\Omega^2_j S_j)}{\mu_j' B_j \mu_j + \text{tr}(\Omega^2_j B_j)} \right]}
\]

(3.6) \quad \text{MSE}_L(\hat{\alpha}_j)

Both formulae yield an approximation to the effect of filtering on the bias, by computing, say, the difference \( \text{Bias}_L(\hat{\alpha}_F) - \text{Bias}_L(\hat{\alpha}_U) \), and on the mean squared error, by the difference \( \text{MSE}_L(\hat{\alpha}_F) - \text{MSE}_L(\hat{\alpha}_U) \). It may be worth emphasizing that so far no restrictions have been imposed on the filter matrix \( F \), nor on the nature of the exogenous regressors. Imposing restrictions on either one or both will greatly simplify the expressions in (3.5) and (3.6). One set of restrictions, arising fairly naturally in the design of filters are those appearing in (2.7). Combined with the fact that the exogenous regressors only contain at most a linear trend and a constant yields the following set of assumptions:

**Assumption 3.1** : Let the filter matrix \( F \) satisfy (2.7) and the regressors be at most a constant and a linear trend, then \( S_j = S, B_j = B \) and \( \mu_j = \mu \) for \( j = U \) or \( F \).

While the assumption is not absolutely necessary to study finite sample behavior, as noted in Section 2, it does make it easier to obtain even simpler formulae. Let us first consider the case where \( \mu = 0 \) followed by that where the mean vector...
is nonzero. The former corresponds, of course, to the AR(1) model without intercept, while the latter is one where a constant and/or a linear trend are included.

**Theorem 3.1**: Let assumption 3.1 hold with $\mu = 0$. Then, the filtering effect on the bias of the estimator $\hat{\alpha}_T$ can be expressed as:

\[
\text{Bias}_L(\hat{\alpha}_T) - \text{Bias}_L(\hat{\alpha}_U) = \\
\frac{(\text{vec}(S'\otimes B'))'\left[(F \otimes U) \otimes (F \otimes U) - (U \otimes F) \otimes (U \otimes F)\right]\text{vec} (\Omega^Y \otimes \Omega^Y)}{(\text{vec}(B' \otimes B'))'\left[(F \otimes U) \otimes (F \otimes U)\right]\text{vec} (\Omega^Y \otimes \Omega^Y)}
\]

Moreover, $E_L(\hat{\alpha}_T)$ is independent of $T$.

**Proof**: See appendix.

Hence, with $\mu = 0$, the Laplace approximation becomes independent of $T$, a consequence of the design of the matrices $S$, $B$ and $\Omega^Y$ in this particular case. As shown in section 4, formula (3.7) can be used to compute the asymptotic bias. This formula is, in fact, relatively simple (one can use small values of $T$) in comparison to those used by Ghysels and Perron (1993) and Hansen and Sargent (1993).

The following result covers the case where $\mu$ is no longer equal to the zero vector, namely:

**Theorem 3.2**: Let Assumption 3.1 hold and let $\mu \neq 0$. Then,

\[
\text{Bias}_L(\hat{\alpha}_T) - \text{Bias}_L(\hat{\alpha}_U) = \frac{\mu'(S - \alpha B)\mu}{\mu B \mu} \left[1 + \frac{(\text{vec} S'_{1})' (F \otimes F) \text{vec} \Omega^Y}{1 + (\text{vec} B'_{1})' (F \otimes F) \text{vec} \Omega^Y}\right] - \\
\frac{1 + (\text{vec} S'_{1})' (U \otimes U) \text{vec} \Omega^Y}{1 + (\text{vec} B'_{1})' (U \otimes U) \text{vec} \Omega^Y}
\]

where $S_{1} = (S - \alpha B) (\mu'(S - \alpha B)\mu)^{-1}$ and $B_{1} = B(\mu' B \mu)^{-1}$.

**Proof**: See appendix.
The result in Theorem 3.2 does not simplify as elegantly as that in Theorem 3.1. Both (3.7) and (3.8) are relatively straightforward to compute, however. Note also that this time the Laplace approximation depends on the sample size.

4. ON THE ADEQUACY OF THE LAPLACE APPROXIMATION

The purpose of this section is to provide an evaluation of the adequacy of the Laplace approximation as a device to study the effect of filtering on dynamic linear regression parameter estimation. The evaluation consists of comparing (1) simulations of bias and mse via Monte Carlo, (2) the Laplace approximation and (3) the asymptotic bias and mse. This will enable us to assess the difference between small sample properties and asymptotic results while at the same time appraise the accurateness of the analytical expression for \( \text{Bias}_L(\hat{\alpha}) \) and \( \text{MSE}_L(\hat{\alpha}) \). We shall consider several situations, all involving the linear approximation to the X-11 filter. In all our experiments for \( \alpha \), we used the grid \( \alpha = -0.95, -0.9, -0.8, \ldots, 0.8, 0.9, 0.95 \). Three different sample sizes were studied, namely \( T + 1 = 60, 120 \) and 180. Since we study the monthly X-11 filter, this consists of 5, 10 and 15 years of data (after filtering). All Monte Carlo simulations involved 10,000 iterations using GAUSS (version 3.0). We now turn to the different cases considered.

(a) The case of \( \mu = 0 \)

The AR(1) model with no intercept is usually taken as a benchmark. It is particularly appropriate to start with this case since, as noted in Theorem 3.1, the Laplace approximation to the bias is independent of \( T \). The Laplace approximation to the MSE, however, does depend on \( T \). We report the results both via graphs and tables. Table 4.1 gives a summary of the differences in bias and MSE. Since the Laplace approximation to the bias of \( \hat{\alpha}_U \) always equals zero, which corresponds to the asymptotically consistent estimator \( \hat{\alpha}_U \), it is not reported in Table 4.1. Moreover, as \( \text{Bias}_L(\hat{\alpha}_U) = 0 \), it follows that the difference in bias simply corresponds to \( \text{Bias}_L(\hat{\alpha}_F) \). The shape of the curve \( \text{Bias}_L(\hat{\alpha}_F) \), which appears as the solid line in Figure 4.2, is rather interesting. First, \( \text{Bias}_L(\hat{\alpha}_F) > 0 \) for all values of \( \alpha \) considered. Second, it has a rotated S-form, with large biases for \( \alpha < -0.8 \) and \( -0.2 < \alpha \leq 0.9 \). By large biases we mean an order of magnitude between 0.02 and 0.05. It should also be noted that the Laplace approximations coincide with the asymptotic biases in the filtered case, reported in Ghysels and Perron (1993, Table 1). Hence, the Laplace approximation in the filtered case is also consistent. Next, we turn to the Monte Carlo simulations.
Table 4.1: Comparison of Laplace Approximation (LAP) and Monte Carlo (MC) Simulated Differences in the Bias and MSE
The Case of $\mu = 0$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>LAP Difference in Bias $T=60$</th>
<th>MC Difference in Bias $T=60$</th>
<th>LAP Difference in Bias $T=120$</th>
<th>MC Difference in Bias $T=120$</th>
<th>LAP Difference in Bias $T=180$</th>
<th>MC Difference in Bias $T=180$</th>
<th>LAP Difference in MSE $T=60$</th>
<th>MC Difference in MSE $T=60$</th>
<th>LAP Difference in MSE $T=120$</th>
<th>MC Difference in MSE $T=120$</th>
<th>LAP Difference in MSE $T=180$</th>
<th>MC Difference in MSE $T=180$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.95</td>
<td>0.0485</td>
<td>-0.0058</td>
<td>0.0178</td>
<td>0.0276</td>
<td>-0.2700</td>
<td>-0.0071</td>
<td>-0.1461</td>
<td>0.005</td>
<td>-0.0993</td>
<td>0.0024</td>
<td>-0.0189</td>
<td>0.0029</td>
</tr>
<tr>
<td>-0.9</td>
<td>0.0326</td>
<td>0.0245</td>
<td>0.0295</td>
<td>0.0309</td>
<td>-0.0674</td>
<td>0.0050</td>
<td>-0.0299</td>
<td>0.0036</td>
<td>-0.0189</td>
<td>0.0029</td>
<td>-0.0189</td>
<td>0.0029</td>
</tr>
<tr>
<td>-0.8</td>
<td>0.0168</td>
<td>0.0252</td>
<td>0.0210</td>
<td>0.0205</td>
<td>0.0054</td>
<td>0.0047</td>
<td>0.0057</td>
<td>0.0023</td>
<td>0.0045</td>
<td>0.0017</td>
<td>0.0061</td>
<td>0.0014</td>
</tr>
<tr>
<td>-0.7</td>
<td>0.0118</td>
<td>0.0217</td>
<td>0.0172</td>
<td>0.0160</td>
<td>0.0132</td>
<td>0.0042</td>
<td>0.0085</td>
<td>0.0022</td>
<td>0.0061</td>
<td>0.0014</td>
<td>0.0061</td>
<td>0.0014</td>
</tr>
<tr>
<td>-0.6</td>
<td>0.0119</td>
<td>0.0205</td>
<td>0.0170</td>
<td>0.0152</td>
<td>0.0114</td>
<td>0.0041</td>
<td>0.0070</td>
<td>0.0024</td>
<td>0.0050</td>
<td>0.0016</td>
<td>0.0050</td>
<td>0.0016</td>
</tr>
<tr>
<td>-0.5</td>
<td>0.0140</td>
<td>0.0222</td>
<td>0.0186</td>
<td>0.0166</td>
<td>0.0087</td>
<td>0.0045</td>
<td>0.0053</td>
<td>0.0027</td>
<td>0.0038</td>
<td>0.0018</td>
<td>0.0038</td>
<td>0.0018</td>
</tr>
<tr>
<td>-0.4</td>
<td>0.0165</td>
<td>0.0216</td>
<td>0.0204</td>
<td>0.0195</td>
<td>0.0066</td>
<td>0.0040</td>
<td>0.0041</td>
<td>0.0027</td>
<td>0.0029</td>
<td>0.0019</td>
<td>0.0029</td>
<td>0.0019</td>
</tr>
<tr>
<td>-0.3</td>
<td>0.0191</td>
<td>0.0230</td>
<td>0.0217</td>
<td>0.0207</td>
<td>0.0053</td>
<td>0.0044</td>
<td>0.0033</td>
<td>0.0029</td>
<td>0.0024</td>
<td>0.0021</td>
<td>0.0024</td>
<td>0.0021</td>
</tr>
<tr>
<td>-0.2</td>
<td>0.0217</td>
<td>0.0242</td>
<td>0.0232</td>
<td>0.0223</td>
<td>0.0045</td>
<td>0.0042</td>
<td>0.0029</td>
<td>0.0027</td>
<td>0.0022</td>
<td>0.002</td>
<td>0.0022</td>
<td>0.002</td>
</tr>
<tr>
<td>-0.1</td>
<td>0.0242</td>
<td>0.0255</td>
<td>0.0249</td>
<td>0.0250</td>
<td>0.0043</td>
<td>0.0042</td>
<td>0.0028</td>
<td>0.0026</td>
<td>0.0021</td>
<td>0.0022</td>
<td>0.0021</td>
<td>0.0022</td>
</tr>
<tr>
<td>0</td>
<td>0.0267</td>
<td>0.0263</td>
<td>0.0264</td>
<td>0.0264</td>
<td>0.0045</td>
<td>0.0038</td>
<td>0.0029</td>
<td>0.0027</td>
<td>0.0022</td>
<td>0.0021</td>
<td>0.0022</td>
<td>0.0021</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0291</td>
<td>0.0264</td>
<td>0.0274</td>
<td>0.0277</td>
<td>0.0051</td>
<td>0.0032</td>
<td>0.0032</td>
<td>0.0026</td>
<td>0.0025</td>
<td>0.0020</td>
<td>0.0025</td>
<td>0.0020</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0315</td>
<td>0.0279</td>
<td>0.0293</td>
<td>0.0303</td>
<td>0.0062</td>
<td>0.0033</td>
<td>0.0038</td>
<td>0.0024</td>
<td>0.0029</td>
<td>0.0020</td>
<td>0.0029</td>
<td>0.0020</td>
</tr>
<tr>
<td>0.3</td>
<td>0.0336</td>
<td>0.0289</td>
<td>0.0300</td>
<td>0.0313</td>
<td>0.0078</td>
<td>0.0027</td>
<td>0.0047</td>
<td>0.0021</td>
<td>0.0036</td>
<td>0.0018</td>
<td>0.0036</td>
<td>0.0018</td>
</tr>
<tr>
<td>0.4</td>
<td>0.0356</td>
<td>0.0293</td>
<td>0.0323</td>
<td>0.0330</td>
<td>0.0101</td>
<td>0.0021</td>
<td>0.0060</td>
<td>0.0017</td>
<td>0.0045</td>
<td>0.0016</td>
<td>0.0045</td>
<td>0.0016</td>
</tr>
<tr>
<td>0.5</td>
<td>0.0371</td>
<td>0.0297</td>
<td>0.0324</td>
<td>0.0336</td>
<td>0.0133</td>
<td>0.0013</td>
<td>0.0077</td>
<td>0.0014</td>
<td>0.0056</td>
<td>0.0013</td>
<td>0.0056</td>
<td>0.0013</td>
</tr>
<tr>
<td>0.6</td>
<td>0.0377</td>
<td>0.0328</td>
<td>0.0340</td>
<td>0.0340</td>
<td>0.0176</td>
<td>0.0003</td>
<td>0.0099</td>
<td>0.0007</td>
<td>0.0071</td>
<td>0.0010</td>
<td>0.0071</td>
<td>0.0010</td>
</tr>
<tr>
<td>0.7</td>
<td>0.0362</td>
<td>0.0309</td>
<td>0.0327</td>
<td>0.0327</td>
<td>0.0229</td>
<td>-0.0007</td>
<td>0.0125</td>
<td>0.0000</td>
<td>0.0088</td>
<td>0.0005</td>
<td>0.0088</td>
<td>0.0005</td>
</tr>
<tr>
<td>0.8</td>
<td>0.0306</td>
<td>0.0219</td>
<td>0.0259</td>
<td>0.0276</td>
<td>0.0276</td>
<td>-0.0015</td>
<td>0.0187</td>
<td>-0.0007</td>
<td>0.0102</td>
<td>-0.0003</td>
<td>0.0102</td>
<td>-0.0003</td>
</tr>
<tr>
<td>0.9</td>
<td>0.0187</td>
<td>0.0122</td>
<td>0.0158</td>
<td>0.0167</td>
<td>0.0291</td>
<td>-0.0013</td>
<td>0.0154</td>
<td>-0.0011</td>
<td>0.0105</td>
<td>-0.0008</td>
<td>0.0105</td>
<td>-0.0008</td>
</tr>
<tr>
<td>0.95</td>
<td>0.0101</td>
<td>0.0052</td>
<td>0.0081</td>
<td>0.0088</td>
<td>0.0266</td>
<td>-0.0001</td>
<td>0.0146</td>
<td>-0.0008</td>
<td>0.0100</td>
<td>-0.0007</td>
<td>0.0100</td>
<td>-0.0007</td>
</tr>
</tbody>
</table>
displayed in Figure 4.1. Here, we note the familiar downward bias of \( \hat{\alpha}_U \) in small samples for positive \( \alpha \). However, we also observe an upward bias for \( \hat{\alpha}_F \), i.e., X-11 filtering reverses the sign of bias over the same range for \( \alpha \). It should be noted that the biases, whether filtered data are used or not, appear quite similar for \( T + 1 = 60, 120 \) and 180.\(^5\)

Let us now turn our attention to the difference in biases, i.e., 
\[
\text{Bias}_L(\hat{\alpha}_F) - \text{Bias}_L(\hat{\alpha}_U),
\]
and compare the analytic results with the Monte Carlo simulations. Such a comparison appears again in Figure 4.2 and is summarized in Table 4.1. Clearly, the Laplace approximation and hence the asymptotic distribution theory in this case are both a very good predictor of the small-sample behavior for differences in bias. The full line in Figure 4.2, which represents the Laplace approximation, tracks the Monte Carlo curves quite closely. The only evidence of discrepancy occurs when \( \alpha \) takes extreme negative values. Since this range of the parameter space is less interesting, this discrepancy is not very important for practical purposes. The numerical results in Table 4.1 report this finding in further detail.

Regarding the MSE, we also find the Laplace approximation to be quite accurate except at the edges of the parameter space, in particular for values near \( -1 \) but also close to one. Near one, however, we find that the approximation error to the effect of filtering on the MSE is still quite acceptable. Figure 4.3 covers the \( T = 60 \) case, while the next figure covers \( T = 180 \). As \( T \) grows, we find much improvement around the extreme negative values for \( \alpha \) and also around one. Overall, the results for the MSE are very satisfactory and quite similar to those of the bias, considering the fact that the region around the extreme negative values is empirically less relevant.

\[(b) \text{ The case of } \mu \neq 0\]

We now turn our attention to a case where \( \mu \neq 0 \). Obviously, many such cases can be constructed. We will focus on a specific one, which is fairly representative, and comment on other particular cases we considered, though do not report all the details. A most relevant case to focus on with \( \mu \neq 0 \) is one with seasonal dummies and a time trend. Because of the presence of seasonal dummies, it should be noted that

\[\text{Similar findings, also based on Monte Carlo simulations, were reported recently in the econometrics textbook by Davidson and MacKinnon (1993).}\]
$X_U$ and $X_F$ are no longer equal. The seasonal pattern chosen has mean shifts equal to
-1.0 for the first three months of the year, 0.0 the next three, 1.0 for the third quarter
months and, finally, 0.0 once again, for the remaining three months. Other patterns of
seasonal level shifts were considered as well, yet this had very little impact on the
results we found. The trend slope coefficient was set to 0.01. Here, we did not find
that the results were robust to perturbations. Indeed, with large trend coefficients, the
Laplace approximation deteriorated considerably, as will be discussed shortly.

The results for the Laplace approximation will again be reported via figures and
details appearing in Table 4.2. This table includes sample sizes $T = 60, 120$ and 180,
while the figures report only the largest and smallest sample size. Figures 4.5 and 4.6
cover $T = 60$, the former displaying the difference in bias obtained via Monte Carlo
and Laplace approximations, followed by the difference in MSE. Hence, we do not
report here the details of the bias itself but only the difference in bias (and MSE)
before and after filtering. Lieberman (1992) provides a detailed discussion of the
Laplace approximation of the bias and MSE in regression models with trends and
intercepts.

Figures 4.5 through 4.8 show patterns that are quite similar to those reported in
the $\mu = 0$ case. Again, for extreme negative values of $\alpha$, the Laplace approximation to
both the bias and MSE tend to lose accuracy. Also, for the MSE approximation, there
is an increase in error as $\alpha$ approaches unity. This error decreases again as $T$ grows. It
should also be noted that as the trend coefficient increases, the accuracy tends to
decrease. Yet, here again one must take extreme values of the trend coefficient like 0.2
or 0.5, which are typically not relevant in most practical circumstances.

5. CONCLUSION

The Laplace method is yet a relatively unexplored tool of econometric theory.
In this paper, we proposed this method to characterize bias in estimation induced by
filtering data series. As many data series in applied econometrics are filtered, such a
tool is quite relevant and useful. The results in the paper can easily be used and
extended for the GLS estimator, since it is assumed that $\Omega$ and $F$ are known so that the
typical GLS transformation can be performed. Other types of approximation methods
could also be explored in further research. Overall, we found that the Laplace
approximation strikes a balance between (1) simplicity of calculations and (2) accuracy
of the approximations. Indeed, the relatively simple formulae seem to describe the
effect of filtering very well both for the bias and MSE, with the exception of values for
$\alpha$ near $-1$, i.e., part of the parameter space which is of lesser practical importance.
Table 4.2: Comparison of Laplace Approximation (LAP) and Monte Carlo Simulated (MC) Differences in the Bias
The Case of $\mu \neq 0$, AR(1) with Trend and Seasonal Dummies

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>(T=60)</th>
<th>(T=120)</th>
<th>(T=180)</th>
<th>(T=60)</th>
<th>(T=120)</th>
<th>(T=180)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LAP</td>
<td>MC</td>
<td>LAP</td>
<td>MC</td>
<td>LAP</td>
<td>MC</td>
</tr>
<tr>
<td>-0.95</td>
<td>0.0150</td>
<td>-0.0557</td>
<td>0.0310</td>
<td>-0.0077</td>
<td>0.0370</td>
<td>0.0111</td>
</tr>
<tr>
<td>-0.9</td>
<td>0.0035</td>
<td>-0.0051</td>
<td>0.0174</td>
<td>0.0119</td>
<td>0.0224</td>
<td>0.0193</td>
</tr>
<tr>
<td>-0.8</td>
<td>0.0068</td>
<td>0.0019</td>
<td>0.0042</td>
<td>0.0093</td>
<td>0.0082</td>
<td>0.0118</td>
</tr>
<tr>
<td>-0.7</td>
<td>0.0083</td>
<td>0.0020</td>
<td>0.0009</td>
<td>0.0070</td>
<td>0.0043</td>
<td>0.0088</td>
</tr>
<tr>
<td>-0.6</td>
<td>0.0059</td>
<td>0.0036</td>
<td>0.0021</td>
<td>0.0074</td>
<td>0.0052</td>
<td>0.0088</td>
</tr>
<tr>
<td>-0.5</td>
<td>0.0026</td>
<td>0.0047</td>
<td>0.0049</td>
<td>0.0090</td>
<td>0.0077</td>
<td>0.0110</td>
</tr>
<tr>
<td>-0.4</td>
<td>0.0004</td>
<td>0.0076</td>
<td>0.0077</td>
<td>0.0116</td>
<td>0.0105</td>
<td>0.0128</td>
</tr>
<tr>
<td>-0.3</td>
<td>0.0027</td>
<td>0.0074</td>
<td>0.0101</td>
<td>0.0135</td>
<td>0.0129</td>
<td>0.0153</td>
</tr>
<tr>
<td>-0.2</td>
<td>0.0042</td>
<td>0.0080</td>
<td>0.0121</td>
<td>0.0142</td>
<td>0.0151</td>
<td>0.0165</td>
</tr>
<tr>
<td>-0.1</td>
<td>0.0051</td>
<td>0.0082</td>
<td>0.0138</td>
<td>0.0144</td>
<td>0.0171</td>
<td>0.0179</td>
</tr>
<tr>
<td>0</td>
<td>0.0056</td>
<td>0.0073</td>
<td>0.0153</td>
<td>0.0154</td>
<td>0.0189</td>
<td>0.0196</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0062</td>
<td>0.0063</td>
<td>0.0168</td>
<td>0.0156</td>
<td>0.0207</td>
<td>0.0201</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0071</td>
<td>0.0060</td>
<td>0.0184</td>
<td>0.0171</td>
<td>0.0226</td>
<td>0.0216</td>
</tr>
<tr>
<td>0.3</td>
<td>0.0086</td>
<td>0.0071</td>
<td>0.0203</td>
<td>0.0180</td>
<td>0.0246</td>
<td>0.0225</td>
</tr>
<tr>
<td>0.4</td>
<td>0.0110</td>
<td>0.0076</td>
<td>0.0226</td>
<td>0.0198</td>
<td>0.0268</td>
<td>0.0245</td>
</tr>
<tr>
<td>0.5</td>
<td>0.0142</td>
<td>0.0093</td>
<td>0.0252</td>
<td>0.0214</td>
<td>0.0291</td>
<td>0.0259</td>
</tr>
<tr>
<td>0.6</td>
<td>0.0181</td>
<td>0.0113</td>
<td>0.0278</td>
<td>0.0232</td>
<td>0.0311</td>
<td>0.0286</td>
</tr>
<tr>
<td>0.7</td>
<td>0.0219</td>
<td>0.0138</td>
<td>0.0293</td>
<td>0.0242</td>
<td>0.0316</td>
<td>0.0281</td>
</tr>
<tr>
<td>0.8</td>
<td>0.0236</td>
<td>0.0141</td>
<td>0.0276</td>
<td>0.0225</td>
<td>0.0287</td>
<td>0.0256</td>
</tr>
<tr>
<td>0.9</td>
<td>0.0211</td>
<td>0.0106</td>
<td>0.0206</td>
<td>0.0162</td>
<td>0.0200</td>
<td>0.0174</td>
</tr>
<tr>
<td>0.95</td>
<td>0.0185</td>
<td>0.0050</td>
<td>0.0151</td>
<td>0.0104</td>
<td>0.0134</td>
<td>0.0109</td>
</tr>
</tbody>
</table>
APPENDIX

Proof of Theorem 3.1: From (3.4) and Assumption 3.1, we know that with \( \mu = 0 \):

\[
\text{Bias}_L(\hat{\alpha}_F) - \text{Bias}_L(\hat{\alpha}_U) = \frac{\text{tr}(\Omega_Y^U(S - \alpha B))}{\text{tr}(\Omega_Y^F B)} - \frac{\text{tr}(\Omega_Y^U(S - \alpha B))}{\text{tr}(\Omega_Y^U B)}
\]

\[
= \frac{\text{tr}(\Omega_Y^U(S - \alpha B)) \text{tr}(\Omega_Y^U B) - \text{tr}(\Omega_Y^U(S - \alpha B)) \text{tr}(\Omega_Y^F B)}{\text{tr}(\Omega_Y^F B) \text{tr}(\Omega_Y^U B)}
\]

Noting that \( \text{tr}(A) \text{tr}(B) = \text{tr}(A \otimes B) \) and \( AC \otimes BD = (A \otimes B)(C \otimes D) \), we find that:

\[
\text{Bias}_L(\hat{\alpha}_F) - \text{Bias}_L(\hat{\alpha}_U) = \frac{\text{tr}(\Omega_Y^U \otimes \Omega_Y^F ((S - \alpha B) \otimes B))}{\text{tr}(\Omega_Y^F \otimes \Omega_Y^U (B \otimes B))} - \frac{\text{tr}(\Omega_Y^U \otimes \Omega_Y^F ((S - \alpha B) \otimes B))}{\text{tr}(\Omega_Y^F \otimes \Omega_Y^U (B \otimes B))}
\]

\[
= \frac{\text{tr}(\Omega_Y^U \otimes \Omega_Y^F (S \otimes B)) - \text{tr}(\Omega_Y^U \otimes \Omega_Y^F (S \otimes B))}{\text{tr}(\Omega_Y^F \otimes \Omega_Y^U (B \otimes B))} - \frac{\alpha}{\text{tr}(\Omega_Y^F \otimes \Omega_Y^U (B \otimes B))}
\]

The last term can be shown to be equal to \( \alpha \), hence canceling out with the second term and yielding:

\[
\text{Bias}_L(\hat{\alpha}_F) - \text{Bias}_L(\hat{\alpha}_U) = \frac{\text{tr}((\Omega_Y^U \otimes \Omega_Y^F) - (\Omega_Y^U \otimes \Omega_Y^F)) (S \otimes B))}{\text{tr}(\Omega_Y^F \otimes \Omega_Y^U (B \otimes B))}.
\]

Next, it must be noted that from (2.7),

\[
\Omega_Y^Y \otimes \Omega_Y^Y = (F \Omega_Y^F) \otimes (F \Omega_Y^U) = (F \otimes U) (\Omega_Y^Y \otimes \Omega_Y^Y) (F \otimes U),
\]

using again \( AC \otimes BD = (A \otimes B)(C \otimes D) \) twice. Moreover, since \( \text{tr}(ABCD) = (\text{vec}D)'(\text{vec}A) \text{vec} B \), from Theorem 3 in Magnus and Neudecker (1988, p. 31), the result in (3.7) follows.
Next, we show that the Laplace approximation is independent of the sample size. When \( \mu = 0 \), 
\[
E_L(\hat{\alpha}_F) = \text{tr}(\Omega^Y_F S) \text{tr}(\Omega^Y_F B). 
\]

The matrices \( S \) and \( B \) with \( \mu = 0 \) are
\[
S = \frac{1}{2}(D_1' D_2 + D_2' D_1) = \frac{1}{2} \begin{bmatrix}
0 & 1 & \cdots & 0 \\
1 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 \\
0 & 0 & \cdots & 1 & 0
\end{bmatrix}
\]
\[
B = D_1' D_1 = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{bmatrix}
\]

Although it is tedious to show, as the filter is symmetric, the matrix \( \Omega^Y_F = FR^\alpha R^\alpha F \) is symmetric and the elements on each diagonal are all the same, giving: 
\[
\text{tr}(\Omega^Y_F S) = \frac{1}{2}(T \Omega^Y_{F21} + T \Omega^Y_{F12}) = T \Omega^Y_{F12}, 
\]
where \( \Omega^Y_{F12} \) is the 1,2 element in \( \Omega^Y_F \).

Further, \( \text{tr}(\Omega^Y_F B) = T \Omega^Y_{F11} \), and hence 
\[
E_L(\hat{\alpha}_F) = \frac{\Omega^Y_{F12}}{\Omega^Y_{F11}}, 
\]
which is independent of \( T \).

Q.E.D.

**Proof of Theorem 3.2**: From (3.8) and Assumption 3.1, we know that with \( \mu \neq 0 \)
\[
\text{Bias}_L(\hat{\alpha}_F) - \text{Bias}_L(\hat{\alpha}_U) = \frac{\mu'(S - \alpha B)\mu}{\mu'B\mu} \left[ 1 + \text{tr}((\mu'(S - \alpha B)\mu)^{-1}\Omega^Y_F(S - \alpha B)) \right] - \frac{1 + \text{tr}((\mu'(S - \alpha B)\mu)^{-1}\Omega^Y_F(S - \alpha B))}{1 + \text{tr}((\mu' B\mu)^{-1}\Omega^Y_U(S - \alpha B))}. 
\]

Setting \( S_1 = (\mu'(S - \alpha B)\mu)^{-1}(S - \alpha B) \) and \( B_1 = (\mu' B\mu)^{-1}B \) as well as steps similar to those in the proof of Theorem 3.1 yields the result appearing in (3.7).

Q.E.D.
REFERENCES


Diciccio, T.J. and A.M. Martin (1991), "Approximations of Marginal Tail Probabilities for a Class of Smooth Functions with Applications to Bayesian and Conditional Inference", *Biometrika* 78, 891–902.


Young, A.H. (1968), "Linear Approximations to the Census and BLS Seasonal Adjustment Methods", *Journal of the American Statistical Association* 63, 445–471.


9207 : Tremblay, Rodrigue, "L’impact fiscal statique et dynamique de l’accession du Québec au statut de pays souverain", 33 pages.


9214 : Mandel, Benedikt, Marc Gaudry and Werner Rothengatter, "Linear or Nonlinear Utility Functions in Logit Models? The Impact of German High Speed Rail Demand Forecasts", 17 pages.


9216 : Canova, Fabio et Éric Ghysels, "Changes in Seasonal Patterns : Are They Cyclical", 38 pages.
9218 : Boismenu, Gérard, Nicolas Gravel et Jean-Guy Loranger, "Régime d'accumulation et approche de la régulation : un modèle à équations simultanées", 27 pages.
9220 : Desruelle, Dominique, Gérard Gaudet et Yves Richelle, "Complementarity, Coordination and Compatibility : An Analysis of the Economics of Systems", 52 pages.
9221 : Desruelle, Dominique et Yves Richelle, "The Investment Dynamics of a Duopoly: The Relative Importance of a Head Start", 34 pages.
9226 : Bonomo, Marco et René Garcia, "Indexation, Staggering and Disinflation", 32 pages.
9232 : Kollmann, Robert, "Consumptions, Real Exchange Rates and the Structure of International Asset Markets", 33 pages.
<table>
<thead>
<tr>
<th>ID</th>
<th>Title</th>
<th>Authors/Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>9302</td>
<td>Lemieux, Thomas, &quot;Unions and Wage Inequality in Canada and in the United States&quot;, 66 pages.</td>
<td></td>
</tr>
<tr>
<td>9303</td>
<td>Lemieux, Thomas, &quot;Estimating the Effects of Unions on Wage Inequality in a Two-Sector Model with Comparative Advantage and Non-Random Selection&quot;, 54 pages.</td>
<td></td>
</tr>
<tr>
<td>9305</td>
<td>Lévy-Garboua, Louis et Claude Montmarquette, &quot;Une étude économétrique de la demande de théâtre sur données individuelles&quot;, 40 pages.</td>
<td></td>
</tr>
<tr>
<td>9306</td>
<td>Montmarquette, Claude, Rachel Houle et Sophie Mahseredjian, &quot;The Determinants of University Dropouts : A Longitudinal Analysis&quot;, 17 pages.</td>
<td></td>
</tr>
<tr>
<td>9307</td>
<td>Gaudry, Marc, Benedikt Mandel et Werner Rothengatter, &quot;A Disaggregate Box-Cox Logit Mode Choice Model of Intercity Passenger Travel in Germany&quot;, 17 pages.</td>
<td></td>
</tr>
<tr>
<td>9311</td>
<td>Gaudry, Marc, &quot;Asymmetric Shape and Variable Tail Thickness in Multinomial Probabilistic Responses to Significant Transport Service Level Changes&quot;, 26 pages.</td>
<td></td>
</tr>
<tr>
<td>9315</td>
<td>Allard, Marie, Camille Bronsard et Lise Salvas-Bronsard, &quot;Ck-Conjugate Expectations and Duality&quot;, 22 pages.</td>
<td></td>
</tr>
<tr>
<td>9317</td>
<td>Bronsard, Camille, &quot;L'histoire de l'économie mathématique racontée à Juliette&quot;, 17 pages.</td>
<td></td>
</tr>
</tbody>
</table>
9321 : Dionne, Georges, Denise Desjardins, Claire Laberge-Nadeau et Urs Magg, "Medical Conditions, Risk Exposure and Truck Drivers' Accidents: an Analysis with Count Data Regression Models", 20 pages.
9324 : Dufour, Jean-Marie et Eric Renault, "Causalités à court et à long terme dans les modèles VAR et ARIMA multivariés", 68 pages.
9327 : Proulx, Pierre-Paul, "Quebec in North America: from a Borderlands to a Borderless Economy: an Examination of its Trade Flows with the U.S.A. at the National and Regional Levels", 24 pages.
9331 : Arcand, Jean-Louis L. et Elise S. Brezis, "Disequilibrium Dynamics During the Great Depression", 64 pages.
9334 : Bonomo, Marco et René Garcia, "Disappointment Aversion as a Solution to the Equity Premium and the Risk-Free Rate Puzzles", 40 pages.