Comparing Nested Predictive Regression Models with Persistent Predictors*

Yan Ge†, Tae-Hwy Lee‡, and Michael W. McCracken§

August 31, 2015
Preliminary

Abstract

Inference on stock return predictability is commonly conducted by the in-sample inference on the coefficient estimator of the predictive regression, for which several problems have been identified such as the finite sample bias (when predictors are weakly stationary) and the non-pivotal and non-standard asymptotic distribution and un-correctable bias (when predictors are persistent), and various solutions to these problems have been suggested. In this paper, we adopt the out-of-sample inference of the predictive regression model by the encompassing statistic (ENC) that was studied by Clark and McCracken (2001) when predictors are weakly stationary. The contribution of this paper is to show that the ENC statistic has the asymptotic standard normal distribution even when predictors are persistent as well as when predictors are weakly stationary. This new result is important for empirical research on stock return predictability. While many technical problems arise for in-sample inference on the predictive regression due to persistence of predictors, the out-of-sample inference based on ENC is actually benefited from persistence of predictors because it makes the super-consistency and the asymptotic normality of the parameter estimation. Monte Carlo simulation shows that the asymptotic results hold in finite samples when predictors are weakly stationary and persistent. An application to the predictive regression of the equity premium reveals strong predictive ability of several persistent predictors.

Key Words: inference on stock return predictability, predictive regression, local to unit root process, out-of-sample inference, encompassing test, asymptotic normality, equity premium.

JEL Classification: C53, E37, E27

*We thank Jihyung Lee, Taesuk Lee, Aman Ullah, and seminar participants at UC Davis, UCLA, UC Riverside, Yonsei University, University of Seoul, Bank of Korea, KDI, ... for useful comments.
†School of Public Finance, Central University of Finance and Economics, Beijing 100081. E-mail: yge001@ucr.edu
‡Department of Economics, University of California, Riverside, CA 92521. E-mail: taelee@ucr.edu
§Research Division, Federal Reserve Bank of St. Louis, MO 63166. Email: michael.w.mccracken@stls.frb.org

Comparing Nested Predictive Regression Models with Persistent Predictors

August 31, 2015
Preliminary

Abstract

Inference on stock return predictability is commonly conducted by the in-sample inference on the coefficient estimator of the predictive regression, for which several problems have been identified such as the finite sample bias (when predictors are weakly stationary) and the non-pivotal and non-standard asymptotic distribution and un-correctable bias (when predictors are persistent), and various solutions to these problems have been suggested. In this paper, we adopt the out-of-sample inference of the predictive regression model by the encompassing statistic (ENC) that was studied by Clark and McCracken (2001) when predictors are weakly stationary. The contribution of this paper is to show that the ENC statistic has the asymptotic standard normal distribution even when predictors are persistent as well as when predictors are weakly stationary. This new result is important for empirical research on stock return predictability. While many technical problems arise for in-sample inference on the predictive regression due to persistence of predictors, the out-of-sample inference based on ENC is actually benefited from persistence of predictors because it makes the super-consistency and the asymptotic normality of the parameter estimation. Monte Carlo simulation shows that the asymptotic results hold in finite samples when predictors are weakly stationary and persistent. An application to the predictive regression of the equity premium reveals strong predictive ability of several persistent predictors.

Key Words: inference on stock return predictability, predictive regression, local to unit root process, out-of-sample inference, encompassing test, asymptotic normality, equity premium.

JEL Classification: C53, E37, E27

1 Introduction

When two non-nested models are compared, Diebold and Mariano (DM 1995) point out that the t-statistic of the mean squared forecast error (MSFE) loss-differential is asymptotically standard normal. However, when two nested models with weakly stationary predictors added to a bigger model are compared, Clark and McCracken (CM 2001, 2005, 2009) point out that the t-statistic of DM behaves quite differently from non-nested case, which may result in non-standard distribution. Also, Clark and West (CW 2006, 2007) point out that due to the finite sample parameter estimation error, the DM statistic tends to be negatively biased under the null hypothesis of the equal predictive ability. The bias-corrected DM statistic can be shown to be the encompassing (ENC thereafter) test of Harvey, Leybourne and Newbold (1998). Under the null hypothesis that a stationary predictor has no predictive power, CM (2001) show that ENC is asymptotically standard normal when the ratio of the number of out-of-sample forecasts (P) over the number of in-sample observations (R) to goes to a finity constant ($P/R \to \pi < \infty$).

The current status of the literature does not deal with the two important cases. One is when $P/R \to \infty$, and the other is when the predictor is persistent. The first is important because some authors recently derived the optimal rolling window width (R) in forecasting under unstable environment. For example, GW (2005) suggested to fix R, namely $R = O\left(T^0\right)$, Inoue and Rossi (2013?) derived the optimal in-sample estimation window width $R = \left(T^{2/3}\right)$, and Sun, Hong, and Wang (2015) derived in-sample estimation window width $R = \left(T^{4/5}\right)$. In these cases when $R = \left(T^{\delta}\right), \delta < 1, P = T + 1 - R = O\left(T\right)$. Hence, $P/R \to \infty$. Another consideration is the power of the test, which depends on P. Therefore, CM (2001) omit the important case when $P/R \to \infty$, and this paper fills the missing part. The second is important as studied by numerous papers on predictive regression such as ...... All of these papers are based on the in-sample inference of the coefficient of the predictor..

Hansen and Timmermann (2015) ... our paper provides an important advantage of using the OOS test. The equivalence may require the weak stationarity of the predictor. (need to check the assumption of HT).

Even in the OOS, CCS will suffer the same problem as it use the persistent predictor to check the correlation between forecast errors and the persistent predictor. The actual advantage of the

ENC comes from the fact that ENC is constructed using only the errors, not predictors.

Phillips IVX papers ...

** Inference on the stock return predictability can be conducted in different ways. A common approach is the *in-sample* inference based on the asymptotic distribution of the coefficient estimator in the predictive regression. Several problems have been identified, such as the finite sample bias (when predictors are stationary) and the non-pivotal and nonstandard asymptotic distribution and uncorrectable bias (when predictors are persistent). Some solutions to these problems have been suggested, e.g., Kothari and Shanken (1997), Campbell and Yogo (2006), Jansson and Moreira (2006), Elliott (2011), Cai and Wang (2014), etc. In particular, Phillips and Magdalinos (2007, 2009) and Phillips and Lee (2013, 2014) show that their IVX approach can produce the asymptotic distribution which is mixed normal. In this paper, we adopt a different approach, the *out-of-sample* inference of the predictive regression model by comparing the forecast error loss as in Diebold-Mariano (DM 1995). Clark and West (2006, 2007) show that DM is biased and under-sized and also show that the bias-corrected DM statistic is equivalent to the encompassing (ENC) statistic. ENC was carefully studied in Clark and McCracken (CM 2001) when predictors are weakly stationary. This paper considers the ENC test with a highly persistent predictor. The under-sized problem of DM is more severe when the predictor is persistent. We show that, unlike the in-sample inference on the predictive regression which is affected by non-stationarity of persistent predictors, the out-of-sample inference based on ENC is robust to near-unit-root non-stationarity of the predictor. In fact, non-stationarity of persistent predictors helps ENC perform even better in finite sample due to the faster rate of convergence in the parameter estimation. We obtain the asymptotic and finite sample properties of the encompassing statistic when predictors has an AR root in the vicinity of or on unity. In contrast to the asymptotic mixed normal distribution or a mixture of normal and nonstandard distribution of the in-sample inference, we show that the out-of-sample inference using ENC has the asymptotic standard normal distribution. An application to the predictive regression of the equity premium reveals strong predictive ability of several persistent predictors.***

In this paper we show that under the null hypothesis, the ENC test restores the asymptotically standard normal when the ratio of the out-of-sample number of forecasts (P) to the in-sample number of observations (R) goes to infinity ($P/R \to \infty$). Under the null hypothesis, the encompassing test also has good power.

CM (2001) and Clark and West (CW 2006, 2007) considered predictive mean regression with weak stationary predictor. This paper considers the predictive mean regression with a highly persistent predictor with an AR root local to unity. We compare two nested regression models using the squared-loss function. We show that DM statistic still tends to be negative under the null hypothesis of the equal predictive ability and is more severely undersized if the predictor is a highly persistent predictor. The t-statistic of encompassing test, in which a positive term is added to correct the negative bias of DM, is a robust test and has the correct size under the null hypothesis. We analytically show that the robustness arises from the super consistency property of the additional predictor from Model 2 that follows Ornstein–Uhlenbeck process, thus the convergent rate of the forecast error from Model 2 is faster than that from Model 1 and the asymptotic distribution of ENC statistic has the same asymptotic distribution as shown in CW (2006, 2007). We use Monte Carlo simulation to compare two different statistics and show that when the highly persistent estimator is added in Model 2, the ENC statistic is robust and has the correct size, whereas DM test is seriously undersized. An application to the predictive regression of the equity premium reveals strong predictive ability of several persistent predictors (such as inflation and interest rate) by ENC, but with little or none can be seen from DM.

The paper is organized as follows. Section 2 illustrates the methods of testing out-of-sample Granger-causality in mean using rolling scheme. Section 3 illustrates the asymptotic distribution of the encompassing test with a weak stationary predictor from CM (2001). Section **??** presents the asymptotic distribution of the encompassing test with a weak stationary predictor when the ratio of out-of-sample to in-sample observation is infinite. Section 4 presents the asymptotic distribution of encompassing test with a highly persistent estimator when the ratio of out-of-sample to in-sample observation is infinite. Section 7 is Monte Carlo simulation to examine the finite sample size and power behavior of the DM and ENC statistics. In Section 8 we present the empirical analysis for Goyal and Welch (2008) in comparing the two nested mean models. Section 9 concludes.

2 Comparing out-of-sample predictive ability of nested models

To test for the out-of-sample predictive ability of x_t for y_{t+1}, we consider the following two nested models with the predictor x_t in Model 2 being local to unit root process:

$$\text{Model 1} \quad : \quad y_{t+1} = x'_{1,t}\beta_{1,t} + e^{(1)}_{t+1} = c_1 + e^{(1)}_{t+1}, \tag{1}$$

$$\text{Model 2} \quad : \quad y_{t+1} = x'_{2,t}\beta_{2,t} + e^{(2)}_{t+1} = c_2 + bx_t + e^{(2)}_{t+1}, \tag{2}$$

where c_i is the constant term for Model i, x_t is the predictor with local to unit autoregressive (AR) root process $x_{t+1} = \phi x_t + v_{t+1}$. We will consider the simple case when $x'_{1,t} = 1$ and $x'_{2,t} = (1\ x_t)'$. Under the null hypothesis, $b = 0$ and $e^{(1)}_{t+1} = e^{(2)}_{t+1}$, denoted as e_{t+1}. At each time t, both c_i and b are estimated with the rolling window of size R up to time t. Therefore

$$\begin{aligned} \hat{c}_{1,t} - c_1 &= B_1(t) H_1(t) \\ \left(\hat{c}_{2,t}, \hat{b}_t\right)' - (c_2, b)' &= B_2(t) H_2(t) \end{aligned}$$

where $x'_{1,t} = 1$, $x'_{2,t} = (1\ x_t)'$, $q_{i,t} = x'_{i,t}x_{i,t}$ for Model i at time t, and $B_i(t) = \left(R^{-1}\sum_{j=t-R}^{t-1} q_{i,j}\right)^{-1}$, $h_{i,t} = x'_{i,t}e_{t+1}$ and $H_i(t) = R^{-1}\sum_{j=t-R}^{t-1} h_{i,t}$. Let $f^{(1)}_{t+1} = \hat{c}_{1,t}$ be the forecasts for Model 1 and $f^{(2)}_{t+1} = \hat{c}_{2,t} + \hat{b}_t x_t$ be the forecast for Model 2 at time t and $\hat{e}^{(1)}_{t+1} = y_{t+1} - f^{(1)}_{t+1}$, $\hat{e}^{(2)}_{t+1} = y_{t+1} - f^{(2)}_{t+1}$ be the forecast errors with the squared forecast-error loss

$$L\left(\hat{e}^{(i)}_{t+1}\right) \equiv \left(\hat{e}^{(i)}_{t+1}\right)^2, \qquad i = 1, 2.$$

To test for equal predictive accuracy of the two models, the null hypothesis is

$$\mathbb{H}_0 : \mathbb{E}\left[L\left(\hat{e}^{(1)}_{t+1}\right) - L\left(\hat{e}^{(2)}_{t+1}\right)\right] = 0. \tag{3}$$

Under $\mathbb{H}_0$, x_t does not Granger-cause y_{t+1} in mean and thus $b = 0$. If x_t Granger-causes y_{t+1}, i.e., $b \neq 0$, thus the alternative hypothesis is

$$\mathbb{H}_1 : \mathbb{E}\left[L\left(\hat{e}^{(1)}_{t+1}\right) - L\left(\hat{e}^{(2)}_{t+1}\right)\right] > 0. \tag{4}$$

The Diebold-Mariano square loss differential is defined as

$$\hat{D}_P = P^{-1}\sum_{t=R}^{T} L\left(\hat{e}^{(1)}_{t+1}\right) - L\left(\hat{e}^{(2)}_{t+1}\right), \tag{5}$$

and the adjusted MSFE loss-differential is defined as

$$\hat{B}_P = P^{-1} \sum_{t=R}^{T} \hat{e}_{t+1}^{(1)} \left(\hat{e}_{t+1}^{(1)} - \hat{e}_{t+1}^{(2)} \right),$$

where R is the number of observations in the rolling windows for the in-sample estimation, P is the number of out-of-sample forecasts, and $R + P = T + 1$. The two statistics are standardized to form the DM statistic $DM_P \equiv \hat{S}_P^{-0.5} \sqrt{P} \hat{D}_P$, and the encompassing statistic $ENC_P \equiv \hat{Q}_P^{-0.5} \sqrt{P} \hat{B}_P$, where $\hat{S}_P$ and $\hat{Q}_P$ are the consistent estimators of $S_P = var\left(\sqrt{P}\hat{D}_P\right)$ and $Q_P = var\left(\sqrt{P}\hat{B}_P\right)$ respectively.

3 Asymptotic distribution of ENC with a stationary predictor (CM 2001)

First, we consider a stationary predictor as in CM (2001, 2005) and CW (2006, 2007).

Assumption 1a. $\{x_t\}$ is a weakly stationary process and $\mathbb{E}(q_{i,t})$ is bounded for all t and $i = 1, 2$. We define $B_i = (\mathbb{E}q_{i,t})^{-1}$ for model $i = 1, 2$.

Let $\pi = \lim_{P,R \to \infty} P/R$ and $\xi = R/T = R/(P + R)$. Note that $1/\xi - 1 \to \pi$. We consider three cases on π :

Assumption 2a. $0 < \pi < \infty$.

Assumption 2b. $\pi = 0$ (or $\xi \to 1$).

Assumption 2c. $\pi = \infty$ (or $\xi \to 0$).

Proposition 1 (CM 2001). Under Assumptions 1a and 2a,

$$ENC_P \Rightarrow \frac{\int_\xi^1 \xi^{-1} \left[W(s) - W(s-\xi)\right] \mathrm{d}W(s)}{\sqrt{\int_\xi^1 \xi^{-2} \left[W(s) - W(s-\xi)\right]^2 \mathrm{d}s}},$$

under $\mathbb{H}_0$, where $W(s)$ is a Wiener process and $s \in [0, 1]$. When Assumption 2a holds, the RHS of Equation (6) is *not* standard normal.

Proposition 2 (CM 2001). Under Assumptions 1a and 2b,

$$ENC_P \Rightarrow \lim_{\xi \to 1} \frac{\int_\xi^1 \xi^{-1} \left[W(s) - W(s-\xi)\right] \mathrm{d}W(s)}{\sqrt{\int_\xi^1 \xi^{-2} \left[W(s) - W(s-\xi)\right]^2 \mathrm{d}s}} \sim N(0, 1), \tag{6}$$

under $\mathbb{H}_0$, where $W(s)$ is a Wiener process and $s \in [0,1]$. When Assumption 2b holds, the RHS of Equation (6) is standard normal.

Remark 1. CM (2001) shows that when Assumption 2b holds ($\xi \to 1, \pi \to 0$) then ENC_P is asymptotically standard normal. However, CM (2001) does not consider the case when Assumption 2c holds ($\xi \to 0, \pi \to \infty$). In Section **??** below, we consider this case and show that ENC_P is still asymptotically standard normal.

Remark 2: CM (2001) assumes Assumption 1a that the predictor $\{x_t\}$ is weakly stationary and shows that ENC_P is asymptotically standard normal under Assumption 2b.

4 Asymptotic distribution of ENC with a persistent predictor when $P/R \to \infty$ (for rolling scheme)

Suppose the predictor x_t in Model 2 follows an AR process $x_{t+1} = \phi x_t + v_{t+1}$ where $\mathbb{E}\left(v_{t+1}^2\right) = \sigma_v^2$. If $|\phi| < 1$, then

$$T^{-1}\sum_{t=1}^{T} x_t^2 \xrightarrow{p} \frac{\sigma_v^2}{1-\phi^2}, \qquad T^{-0.5}\sum_{t=1}^{T} x_t v_{t+1} \Rightarrow N\left(0, \frac{\sigma_v^2}{1-\phi^2}\right),$$

as $T \to \infty$. Many recent papers generalize the above to the case when ϕ approaches to 1 as the sample size T increases, see Bobkoski (1983), Cavanagh (1985), Chan and Wei (1987), Giraitis and Phillips (2006), Mikusheva (2007, 2015), Park (2003), Phillips (1987), Phillips and Lee (2013), and Stock (1991). Let $\phi = 1 - c/T$ for some fixed constant $c \geq 0$, $t = [Tr]$, $r \in [0\ 1]$. Let $x_{[Tr]}/\sqrt{T} \Rightarrow J_x^c(r) = \int_0^r e^{(r-s)c} dB_x(s)$ be an Onstein-Uhlenbeck process and B_x is a Brownian motion. If the AR coefficient ϕ is local to unity, then

$$T^{-2}\sum_{t=1}^{T} x_t^2 \Rightarrow \int_0^1 J_x^c(r)^2\, \mathrm{d}r, \qquad T^{-1}\sum_{t=1}^{T} x_t v_{t+1} \Rightarrow \int_0^1 J_x^c(r)\, \mathrm{d}B_x(r),$$

as $T \to \infty$. To consider the persistent predictor we take the local to unit root process in the following Assumption 1b.

Assumption 1b. $\{x_t\}$ follows an AR process with a root local to unity, $\phi = 1 - c/T$, for some fixed constant $c \geq 0$.

Let $t \equiv [Ts]$ and $\xi \equiv R/T$. Then we have $t/T \to s$ and $(t - R + 1)/T \to (s - \xi)$.

Under Assumption 1b,

$$T^{-2}\sum_{j=t-R+1}^{t}x_j^2\Rightarrow\int_{s-\xi}^{s}J_x^c\left(r\right)^2\mathrm{d}r,\qquad T^{-1}\sum_{j=t-R+1}^{t}x_jv_{j+1}\Rightarrow\int_{s-\xi}^{s}J_x^c\left(r\right)\mathrm{d}B_x\left(r\right),\qquad t=R,\ldots,T,$$

as $T\to\infty$. Now, we state the main result, for the numerator of ENC_P, that is $\sqrt{P}\hat{B}_P$.

Proposition 3. Under Assumptions 1b and 2c, we have

$$\sum\nolimits_{t=R}^{T}\hat{e}_{t+1}^{(1)}\left(\hat{e}_{t+1}^{(1)}-\hat{e}_{t+1}^{(2)}\right)=-\sum\nolimits_{t=R}^{T}e_{t+1}\left(e_{t+1}-\hat{e}_{t+1}^{(1)}\right)+o\left(\xi^{-1}\right)$$

under $\mathbb{H}_0$.

Proof: Under the null hypothesis that $b=0$, $e_{t+1}^{(1)}=e_{t+1}^{(2)}=:e_{t+1}$. Note that

$$\hat{e}_{t+1}^{(i)}=e_{t+1}-x_{i,t}'\left(\hat{\beta}_{i,t}-\beta_i\right)$$

for Model i. Recall $x_{1,t}'=1$. For $\hat{B}_P$, the numerator of ENC_P, we decompose

$$\begin{aligned}
&\sum\nolimits_{t=R}^{T}\hat{e}_{t+1}^{(1)}\left(\hat{e}_{t+1}^{(1)}-\hat{e}_{t+1}^{(2)}\right)\\
=\;&\sum\nolimits_{t=R}^{T}\left[e_{t+1}-x_{1,t}'\left(\hat{\beta}_{1,t}-\beta_{1,t}\right)\right]\left(e_{t+1}-x_{1,t}'\left(\hat{\beta}_{1,t}-\beta_1\right)-e_{t+1}+x_{2,t}'\left(\hat{\beta}_{2,t}-\beta_2\right)\right)\\
=\;&\sum\nolimits_{t=R}^{T}\left[e_{t+1}-x_{1,t}'\left(\hat{\beta}_{1,t}-\beta_1\right)\right]\left(-x_{1,t}'\left(\hat{\beta}_{1,t}-\beta_1\right)+x_{2,t}'\left(\hat{\beta}_{2,t}-\beta_2\right)\right)\\
=\;&\sum\nolimits_{t=R}^{T}e_{t+1}\left[-x_{1,t}'\left(\hat{\beta}_{1,t}-\beta_1\right)\right]+\sum\nolimits_{t=R}^{T}e_{t+1}\left[x_{2,t}'\left(\hat{\beta}_{2,t}-\beta_2\right)\right]\\
&+\sum\nolimits_{t=R}^{T}\left(\hat{\beta}_{1,t}-\beta_1\right)x_{1,t}x_{1,t}'\left(\hat{\beta}_{1,t}-\beta_1\right)-\sum\nolimits_{t=R}^{T}\left(\hat{\beta}_{1,t}-\beta_2\right)x_{1,t}x_{2,t}'\left(\hat{\beta}_{2,t}-\beta_2\right)\\
\equiv\;&A_1+A_2+A_3+A_4
\end{aligned}\tag{7}$$

Lemmas 1-3 show that $A_1+A_2+(A_3+A_4)=O\left(\frac{T}{R}\right)+O\left(\frac{P}{T}\right)+o\left(1\right)$. Hence (7) is dominated by A_1 because $\frac{T}{R}\to\infty$ and $\frac{P}{T}\to1$ under Assumption 2c. □

Lemma 1. Under Assumptions 1b and 2c, $A_1\Rightarrow-\sigma_e^2\xi^{-1}\int_{\xi}^{1}[W(s)-W(s-\xi)]\mathrm{d}W(s)=O\left(\xi^{-1}\right)=O\left(\frac{T}{R}\right)$ under $\mathbb{H}_0$.

Proof: Following Lemma A6 of CM (2001), we show

$$\begin{aligned} A_1 &= \sum\nolimits_{t=R}^{T} e_{t+1}\left[-x_{1,t}'\left(\hat{\beta}_{1,t}-\beta_1\right)\right] \\ &= -\sum\nolimits_{t=R}^{T} e_{t+1}\left(e_{t+1}-\hat{e}_{t+1}^{(1)}\right) \\ &= -\sum\nolimits_{t=R}^{T} e_{t+1}\left(R^{-1}\sum\nolimits_{j=t-R+1}^{t} e_j\right) \\ &= -\sum\nolimits_{t=R}^{T} e_{t+1}\left(T^{-1}\sum\nolimits_{j=t-R+1}^{t} e_j\right)/\xi \\ &= -\sum\nolimits_{t=R}^{T}\left[\left(T^{-1/2}e_{t+1}\right)\left(T^{-1/2}\sum\nolimits_{j=1}^{t} e_{,j}-T^{-1/2}\sum\nolimits_{j=1}^{t-R} e_{,j}\right)\right]/\xi \\ &\Longrightarrow -\sigma_e^2\xi^{-1}\int_{\xi}^{1}\left[W(s)-W(s-\xi)\right]\mathrm{d}W(s). \end{aligned}$$

□

Lemma 2. Under Assumptions 1b and 2c, $A_2 = \sum\nolimits_{t=R}^{T} e_{t+1}\left[x_{2,t}'\left(\hat{\beta}_{2,t}-\beta_{2,t}\right)\right]$ is $O(1-\xi) = O\left(\frac{P}{T}\right)$ under $\mathbb{H}_0$.

Proof: Rewrite

$$\begin{aligned} A_2 &= \sum\nolimits_{t=R}^{T} e_{t+1}x_{2,t}'\left(\hat{\beta}_{2,t}-\beta_2\right) \\ &= \sum\nolimits_{t=R}^{T} e_{t+1}x_{2,t}'\left(\sum_{j=t-R}^{t-1} x_{2,j}x_{2,j}'\right)^{-1}\left(\sum_{j=t-R}^{t-1} x_{2,j}e_{j+1}\right) \\ &= \sum\nolimits_{t=R}^{T} e_{t+1}x_{2,t}'G_T^{-1}\left[G_T^{-1}\sum\nolimits_{j=t-R+1}^{t} x_{2,j}x_{2,j}'G_T^{-1}/\xi\right]^{-1}\left[G_T^{-1}\sum\nolimits_{j=t-R+1}^{t} x_{2,j}e_{j+1}/\xi\right], \end{aligned}$$

where $G_T = \mathrm{diag}\left(T^{0.5}, T\right)$ as before, and for the two bracketed terms in the last line, we have

$$\begin{aligned} G_T^{-1}\left(\sum\nolimits_{j=t-R+1}^{t} x_{2,j}x_{2,j}'\right)G_T^{-1}/\xi &\Rightarrow \begin{pmatrix} \xi & \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}r \\ \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}r & \int_{s-\xi}^{s} \left(J_x^c(r)\right)^2\mathrm{d}r \end{pmatrix}/\xi \sim O(1), \\ G_T^{-1}\sum\nolimits_{j=t-R+1}^{t} x_{2,j}e_{j+1}/\xi &\Rightarrow \begin{pmatrix} \int_{s-\xi}^{s} 1\mathrm{d}W(r) \\ \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}W(r) \end{pmatrix}/\xi \sim O(1), \end{aligned}$$

where $J_x^c(r)$ is an Ornstein-Uhlenbeck process, and $W(r)$ is a Wiener process. Hence

$$\begin{aligned} A_2 &= \sum_{t=R}^{T} e_{t+1} x_{2,t}' G_T^{-1} \left[G_T^{-1} \sum_{j=t-R+1}^{t} x_{2,j} x_{2,j}' G_T^{-1} / \xi \right]^{-1} \left[G_T^{-1} \sum_{j=t-R+1}^{t} x_{2,j} e_{j+1} / \xi \right] \\ &\Rightarrow \int_{\xi}^{1} \begin{pmatrix} 1 & J_x^c(s) \end{pmatrix} \begin{pmatrix} \xi & \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}r \\ \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}r & \int_{s-\xi}^{s} J_x^c(r)^2\,\mathrm{d}r \end{pmatrix}^{-1} \begin{pmatrix} \int_{s-\xi}^{s} 1\mathrm{d}W(r) \\ \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}W(r) \end{pmatrix} \mathrm{d}W(s) \\ &= \int_{\xi}^{1} \begin{pmatrix} 1 & J_x^c(s) \end{pmatrix} \begin{pmatrix} O(\xi) & O(\xi) \\ O(\xi) & O(\xi) \end{pmatrix}^{-1} \begin{pmatrix} O(\xi) \\ O(\xi) \end{pmatrix} \mathrm{d}W(s) \\ &= O(1-\xi) \end{aligned}$$

Therefore $A_2 = \sum_{t=R}^{T} e_{t+1} x_{2,t}' \left(\hat{\beta}_{2,t} - \beta_{2,t} \right) = O(1-\xi)$. □

Lemma 3. Under Assumptions 1b and 2c, $A_3 + A_4$ is $o(1)$ under $\mathbb{H}_0$.

Proof: Let $E_T = \operatorname{diag}(T^0, T^{0.5})$, $F_T = \operatorname{diag}(T^1, T^{1.5})$, $G_T = \operatorname{diag}(T^{0.5}, T^1)$, then for any 2×2 matrix K, we have $E_T F_T = G_T G_T$ and

$$E_T \times K \times F_T = G_T \times K \times G_T,$$

because E_T, F_T, G_T are diagonal. Therefore

$$\begin{aligned} A_3 + A_4 &= \sum_{t=R}^{T} \left(\hat{\beta}_{1,t} - \beta_{1,t} \right) x_{1,t} x_{1,t}' \left(\hat{\beta}_{1,t} - \beta_{1,t} \right) - \sum_{t=R}^{T} \left(\hat{\beta}_{1,t} - \beta_{1,t} \right) x_{1,t} x_{2,t}' \left(\hat{\beta}_{2,t} - \beta_{2,t} \right) \\ &= \sum_{t=R}^{T} H_1'(t) B_1(t) q_{1,t} B_1(t) H_1(t) - \sum_{t=R}^{T} H_1'(t) B_1(t) x_{1,t} x_{2,t}' B_2(t) H_2(t), \end{aligned}$$

where the second line appears to be the same as the second bracketed right-hand side term in (A7) of Lemma A10 in CM (2001), which shows that the above is $o(1)$ under Assumption 1a. However, under Assumption 1b, x_t has an AR root local to unity. We show below that the local-to-unit root in x does not affect Lemma A10 of CM (2001). This is because terms involving x can be suitably normalized as follows

$$\begin{aligned} A_3 + A_4 &= \sum_{t=R}^{T} H_1'(t) B_1(t) q_{1,t} B_1(t) H_1(t) \\ &\quad - \sum_{t=R}^{T} H_1'(t) B_1(t) x_{1,t} \left(x_{2,t}' E_T^{-1} \right) \left[E_T \times R^{-1} B_2(t) \times F_T \times \xi \right] \left[F_T^{-1} \times R H_2(t) / \xi \right] \\ &\equiv \sum_{t=R}^{T} H_1'(t) B_1(t) q_{1,t} B_1(t) H_1(t) - \sum_{t=R}^{T} H_1'(t) B_1(t) x_{1,t} \ddot{x}_{2,t}' \ddot{B}_2(t) \ddot{H}_2(t). \end{aligned}$$

where

$$\ddot{x}_{2,t}' \equiv x_{2,t}' E_T^{-1} \Rightarrow \begin{pmatrix} 1 & J_x^c(r) \end{pmatrix} = O(1),$$

$$
\begin{aligned}
\ddot{B}_2(t) &\equiv E_T \times R^{-1}B_2(t) \times F_T \times \xi \\
&= G_T \left[R^{-1}B_2(t)\right] G_T \times \xi \\
&= \left[G_T^{-1} \left[R^{-1}B_2(t)\right]^{-1} G_T^{-1}\right]^{-1} \times \xi \\
&= \left[G_T^{-1} \sum\nolimits_{j=t-R+1}^{t} x_{2,j}x_{2,j}' G_T^{-1}\right]^{-1} \times \xi \\
&\Rightarrow \begin{pmatrix} \xi & \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}r \\ \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}r & \int_{s-\xi}^{s} \left(J_x^c(r)\right)^2 \mathrm{d}r \end{pmatrix}^{-1} \times \xi \\
&= \begin{pmatrix} O(\xi) & O(\xi) \\ O(\xi) & O(\xi) \end{pmatrix}^{-1} \times O(\xi) = O(1),
\end{aligned}
$$

and

$$
\begin{aligned}
\ddot{H}_2(t) &\equiv F_T^{-1} \times RH_2(t)/\xi \\
&= F_T^{-1} \sum\nolimits_{j=t-R+1}^{t} x_{2,j}e_{j+1}/\xi \\
&= \begin{pmatrix} T^{-1}/\xi \times \sum\nolimits_{j=t-R+1}^{t} e_{j+1} \\ T^{-1.5}/\xi \times \sum\nolimits_{j=t-R+1}^{t} x_j e_{j+1} \end{pmatrix} \\
&= \begin{pmatrix} T^{-0.5}/\xi \times T^{-0.5} \sum\nolimits_{j=t-R+1}^{t} e_{j+1} \\ T^{-0.5}/\xi \times T^{-1} \sum\nolimits_{j=t-R+1}^{t} x_j e_{j+1} \end{pmatrix} \\
&\Rightarrow \begin{pmatrix} T^{-0.5}/\xi \times \int_{s-\xi}^{s} 1\mathrm{d}W(r) \\ T^{-0.5}/\xi \times \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}W(r) \end{pmatrix} \\
&= \begin{pmatrix} O\left(T^{-0.5}/\xi\right) \times O(\xi) \\ O\left(T^{-0.5}/\xi\right) \times O(\xi) \end{pmatrix} = O\left(T^{-0.5}\right).
\end{aligned}
$$

Therefore, $\ddot{x}_{2,t}', \ddot{B}_2(t), \ddot{H}_2(t)$ have the same orders of magnitude as $x_{2,t}', B_2(t), H_2(t)$ in stationary case of Lemma A10 in CM (2001). Therefore $A_3 + A_4$ is $o(1)$ not only under Assumption 1a but also under Assumption 1b. □

Based on Lemmas 1-3 under Assumption 1b and Assumption 2c, $\sum_{t=R}^{T} \hat{e}_{t+1}^{(1)} \left(\hat{e}_{t+1}^{(1)} - \hat{e}_{t+1}^{(2)}\right) = -\sum_{t=R}^{T} e_{t+1} \left(e_{t+1} - \hat{e}_{t+1}^{(1)}\right) + o(1)$. Hence, this is the encompassing test for the martingale difference model $y_{t+1} = e_{t+1}$ and the constant mean model $y_{t+1} = c + e_{t+1}^{(1)}$, as studied by CW (2006).

Proposition 4. Under Assumptions 1b and 2c, the asymptotic distribution of ENC_P is

$$\frac{-\sigma_e^2 \xi^{-1} \int_\xi^1 \left[W(s) - W(s-\xi)\right] \mathrm{d}W(s)}{\sqrt{\sigma_e^4 \times \xi^{-2} \int_\xi^1 \left[W(s) - W(s-\xi)\right]^2 \mathrm{d}s}}$$

under $\mathbb{H}_0$.

Proof: From Lemma 1, $A_1 = O\left(\xi^{-1}\right)$ is the dominant term of in $\hat{B}_P$ and hence ENC_P is

$$\begin{aligned} ENC_P &= A_1/\sqrt{\mathrm{var}\left(A_1\right)} + o\left(1\right) \\ &= \frac{-\sum_{t=R}^{T} e_{t+1}\left(e_{t+1} - \hat{e}_{t+1}^{(1)}\right)}{\sqrt{\sum_{t=R}^{T}\left[-e_{t+1}\left(e_{t+1} - \hat{e}_{t+1}^{(1)}\right) - \hat{c}_P\right]^2}} + o\left(1\right) \\ &\Rightarrow \lim_{\xi \to 0} \frac{-\sigma_e^2 \xi^{-1} \int_\xi^1 \left[W(s) - W(s-\xi)\right] \mathrm{d}W(s)}{\sqrt{\sigma_e^4 \times \xi^{-2} \int_\xi^1 \left[W(s) - W(s-\xi)\right]^2 \mathrm{d}s}} \end{aligned}$$

where $A_1 \Rightarrow -\sigma_e^2 \xi^{-1} \int_\xi^1 \left[W(s) - W(s-\xi)\right] \mathrm{d}W(s)$, $c_{t+1} = -e_{t+1}\left(e_{t+1} - \hat{e}_{t+1}^{(1)}\right)$ and $\hat{c}_P = P^{-1} \sum_{t=R}^{T} c_{t+1}$ $= P^{-1} A_1$. The denominator follows from Lemma 4. □

Lemma 4. Under Assumptions 1b and 2c,

$$\sum_{t=R}^{T}\left[-e_{t+1}\left(e_{t+1} - \hat{e}_{t+1}^{(1)}\right) - \hat{c}_P\right]^2 \Rightarrow \sigma_e^4 \times \xi^{-2} \int_\xi^1 \left[W(s) - W(s-\xi)\right]^2 ds.$$

Proof: Following Lemma A11 of CM (2001), we have

$$\begin{aligned} & \sum_{t=R}^{T}\left[-e_{t+1}\left(e_{t+1} - \hat{e}_{t+1}^{(1)}\right) - \hat{c}_P\right]^2 \\ = & \sum_{t=R}^{T}\left[e_{t+1}\left(e_{t+1} - \hat{e}_{t+1}^{(1)}\right)\right]^2 - P\hat{c}_P^2 \\ = & \sum_{t=R}^{T}\left[e_{t+1}\left(R^{-1}\sum_{j=t-R+1}^{t} e_j\right)\right]^2 + O\left(P^{-1} \times \xi^{-2}\right) \\ = & \frac{T^2}{R^2}\sum_{t=R}^{T}\left[\left(e_{t+1}\right)^2\left(T^{-1/2}\sum_{j=1}^{t} e_{,j} - T^{-1/2}\sum_{j=1}^{t-R} e_{,j}\right)^2\right]\frac{1}{T} + O\left(P^{-1} \times \xi^{-2}\right) \\ \Rightarrow & \ \xi^{-2} \int_\xi^1 \sigma_e^2 \left[\sigma_e W(s) - \sigma_e W(s-\xi)\right]^2 \mathrm{d}s, \end{aligned}$$

where line 3 follows from Lemma 1 for $P\hat{c}_P = A_1 = O\left(\xi^{-1}\right)$. □

Next, we show that under Assumptions 1a and 2c, under $\mathbb{H}_0$,

$$\frac{-\sigma_e^2 \xi^{-1} \int_\xi^1 \left[W(s) - W(s-\xi)\right] \mathrm{d}W(s)}{\sqrt{\sigma_e^4 \times \xi^{-2} \int_\xi^1 \left[W(s) - W(s-\xi)\right]^2 \mathrm{d}s}} \sim N\left(0, 1\right)$$

Proposition 5.

$$\lim_{\psi\to 0}\frac{\int_{\psi}^{1}\psi^{-1}\left[W(s)-W(s-\psi)\right]\mathrm{d}W(s)}{\sqrt{\int_{\psi}^{1}\psi^{-2}\left[W(s)-W(s-\psi)\right]^{2}\mathrm{d}s}}\sim N\left(0,1\right), \tag{8}$$

Proof: We firstly consider the numerator of equation (8) by dividing $[0,1]$ to n equal segments and let $t=[Ts]$, where $[Ts]$ is the integer part of Ts and $s\in[0,1]$. Since ψ is sufficiently small, we can write $\psi\equiv 1/n$. We discretize both the numerator and the denominator. Let $\{u_i\}_{i=1}^{n}$ be a mixing sequence drawn from the standard normal distribution $N(0,1)$ with $E(u)=0$ and $\mathrm{var}(u)=1$. Let $U_t=\sum_{i=1}^{t}u_i$ be the partial sum. Then we have $U_t=\sum_{i=1}^{t}u_i\sim N(0,t)$ and therefore

$$\frac{U_t}{\sqrt{n}}=\frac{\sum_{i=1}^{t}u_i}{\sqrt{n}}\equiv U_n\left(s\right)\Rightarrow W(s),$$

where $U_n(s)$ is a '*cadlag*' function and $W(s)$ is a Wiener process. Note that

$$\begin{aligned} n^{-1}\sum_{t=1}^{n}u_{t-1}u_t &= n^{-1}\sum_{t=1}^{n}U_{t-1}u_t-n^{-1}\sum_{t=1}^{n}U_{t-2}u_t\\ &\Rightarrow \int_{\psi}^{1}W(s)\mathrm{d}W(s)-\int_{\psi}^{1}W(s-\psi)\mathrm{d}W(s)\\ &= \int_{\psi}^{1}\left[W(s)-W(s-\psi)\right]\mathrm{d}W(s)\end{aligned}$$

Considering the term $\int_{\psi}^{1}\left[W(s)-W(s-\psi)\right]^{2}\mathrm{d}s$ in the denominator, we have

$$n^{-2}\sum_{t=1}^{n}u_{t-1}^{2}=n^{-2}\sum_{t=1}^{n}\left(U_{t-1}-U_{t-2}\right)^{2}\Rightarrow\int_{\psi}^{1}\left[W(s)-W(s-\psi)\right]^{2}\mathrm{d}S.$$

We construct the an AR(1) regression model, regressing $\{u_{t+1}\}$ on $\{u_t\}$:

$$u_{t+1}=\delta u_t+e_t$$

The estimator $\hat{\delta}$ equals $\left(\sum_{t=1}^{n}u_{t-1}u_t\right)/\left(\sum_{t=1}^{n}u_{t-1}^{2}\right)$ and the variance $\hat{\delta}$ equals $\left(\sum_{t=1}^{n}u_{t-1}^{2}\right)^{-1}\mathrm{var}(u)=\left(\sum_{t=1}^{n}u_{t-1}^{2}\right)^{-1}$. Therefore Equation (8) can be approximated by

$$\frac{\int_{\psi}^{1}\psi^{-1}\left[W(s)-W(s-\psi)\right]\mathrm{d}W(s)}{\sqrt{\psi^{-2}\int_{\psi}^{1}\left[W(s)-W(s-\psi)\right]^{2}\mathrm{d}s}}\Rightarrow\frac{\sum_{t=1}^{n}u_{t-1}u_t}{\sqrt{\sum_{t=1}^{n}u_{t-1}^{2}}}\sim N(0,1).$$

□

Remark 3: We have consider the nested models in which the null model contains a constant term need to be estimated. We now consider the nested model analogue to Clark and West (2006) in which the null model does not contain a constant term and the error term is martingale difference series

$$\text{Model 1} \quad : \quad y_{t+1} = 0 + e_{t+1}^{(1)}, \tag{9}$$

$$\text{Model 2} \quad : \quad y_{t+1} = x_{2,t}'\beta_{2,t} + e_{t+1}^{(2)} = c + bx_t + e_{t+1}^{(2)}, \tag{10}$$

hence in null model we impose 0 as predictors and the forecast error is the true error term. therefore $\hat{e}_{t+1}^{(1)} = y_{t+1} = e_{t+1}^{(1)}$, $\hat{e}_{t+1}^{(2)} = y_{t+1} - f_{t+1}^{(2)}$. We have the following result (proposition 6) when $c_1 = 0$.

Proposition 6. Under Assumptions 1b and 2c, if Model 1 does not involve any parameter estimation, then $\lim_{\xi \to 0} ENC_P \Rightarrow N(0,1)$ under $\mathbb{H}_0$.

Proof: Under the null hypothesis that $\{x_{1,t}\} \equiv 0$ Therefore Equation (7) only has A_2 term, whose limiting distribution is

$$\int_{\xi}^{1} \begin{pmatrix} 1 & J_x^c(s) \end{pmatrix} \begin{pmatrix} \xi & \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}r \\ \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}r & \int_{s-\xi}^{s} J_x^c(r)^2\,\mathrm{d}r \end{pmatrix}^{-1} \begin{pmatrix} \int_{s-\xi}^{s} 1\mathrm{d}W(r) \\ \int_{s-\xi}^{s} J_x^c(r)\,\mathrm{d}W(r) \end{pmatrix} \mathrm{d}W(s)$$

In the next section, Monte Carlo simulation shows that ENC test is standard normal. *So, we know that the above expression is standard normal! But we still need to prove it!!*

5 Asymptotic distribution of ENC with a persistent predictor when $P/R \to \infty$ (for recursive scheme)

We denote $x_{1,t}' = 1$, $x_{2,t}' = (1\ x_t)'$, $q_{i,t} = x_{i,t}'x_{i,t}$ for Model i at time t, and $B_i(t) = \left(t^{-1}\sum_{j=1}^{t} q_{i,j}\right)^{-1}$, $h_{i,t} = x_{i,t}'e_{t+1}$ and $H_i(t) = t^{-1}\sum_{j=1}^{t} h_{i,t}$. Under Assumption 1b,

$$\begin{aligned} T^{-2}\sum_{j=1}^{t} x_j^2 &\Rightarrow \int_0^s J_x^c(r)^2\,\mathrm{d}r \\ T^{-1}\sum_{j=1}^{t} x_j v_{j+1} &\Rightarrow \int_0^s J_x^c(r)\,\mathrm{d}B_x(r), \qquad t = R, \ldots, T, \end{aligned}$$

as $T \to \infty$. Now, we state the main result, for the numerator of ENC_P, that is $\sqrt{P}\hat{B}_P$.

Proposition 7. Under Assumptions 1b and 2c, we have

$$\begin{aligned} & \sum_{t=R}^{T} \hat{e}_{t+1}^{(1)} \left(\hat{e}_{t+1}^{(1)} - \hat{e}_{t+1}^{(2)} \right) \\ = & \sum_{t=R}^{T} e_{t+1} \left[-x_{1,t}' \left(\hat{\beta}_{1,t} - \beta_1 \right) \right] + \sum_{t=R}^{T} e_{t+1} \left[x_{2,t}' \left(\hat{\beta}_{2,t} - \beta_2 \right) \right] + o_p(1) \end{aligned}$$

under $\mathbb{H}_0$.

Proof: Under the null hypothesis that $b = 0$, $e_{t+1}^{(1)} = e_{t+1}^{(2)} =: e_{t+1}$. Note that

$$\hat{e}_{t+1}^{(i)} = e_{t+1} - x_{i,t}' \left(\hat{\beta}_{i,t} - \beta_i \right)$$

for Model i. Recall $x_{1,t}' = 1$. For $\hat{B}_P$, the numerator of ENC_P, we decompose

$$\begin{aligned} & \sum_{t=R}^{T} \hat{e}_{t+1}^{(1)} \left(\hat{e}_{t+1}^{(1)} - \hat{e}_{t+1}^{(2)} \right) \\ = & \sum_{t=R}^{T} \left[e_{t+1} - x_{1,t}' \left(\hat{\beta}_{1,t} - \beta_{1,t} \right) \right] \left(e_{t+1} - x_{1,t}' \left(\hat{\beta}_{1,t} - \beta_1 \right) - e_{t+1} + x_{2,t}' \left(\hat{\beta}_{2,t} - \beta_2 \right) \right) \\ = & \sum_{t=R}^{T} \left[e_{t+1} - x_{1,t}' \left(\hat{\beta}_{1,t} - \beta_1 \right) \right] \left(-x_{1,t}' \left(\hat{\beta}_{1,t} - \beta_1 \right) + x_{2,t}' \left(\hat{\beta}_{2,t} - \beta_2 \right) \right) \\ = & \sum_{t=R}^{T} e_{t+1} \left[-x_{1,t}' \left(\hat{\beta}_{1,t} - \beta_1 \right) \right] + \sum_{t=R}^{T} e_{t+1} \left[x_{2,t}' \left(\hat{\beta}_{2,t} - \beta_2 \right) \right] \\ & + \sum_{t=R}^{T} \left(\hat{\beta}_{1,t} - \beta_1 \right) x_{1,t} x_{1,t}' \left(\hat{\beta}_{1,t} - \beta_1 \right) - \sum_{t=R}^{T} \left(\hat{\beta}_{1,t} - \beta_2 \right) x_{1,t} x_{2,t}' \left(\hat{\beta}_{2,t} - \beta_2 \right) \\ \equiv & A_1 + A_2 + A_3 + A_4 \end{aligned} \tag{11}$$

Lemmas 5-7 show that $A_1 + A_2 + (A_3 + A_4) = A_1 + A_2 + o_p(1)$. Hence (7) is dominated by $A_1 + A_2$ under Assumption 2c. □

Lemma 5. Under Assumptions 1b and 2c, $A_1 \Rightarrow -\sigma_e^2 \int_0^1 s^{-1} W(s) \mathrm{d}W(s)$. under $\mathbb{H}_0$.

Proof: Following Lemma A6 of CM (2001), we show

$$\begin{aligned} A_1 &= \sum_{t=R}^{T} e_{t+1} \left[-x_{1,t}' \left(\hat{\beta}_{1,t} - \beta_1 \right) \right] \\ &= -\sum_{t=R}^{T} e_{t+1} \left(e_{t+1} - \hat{e}_{t+1}^{(1)} \right) \\ &= -\sum_{t=R}^{T} e_{t+1} \left(t^{-1} \sum_{j=1}^{t} e_j \right) \\ &= -\sum_{t=R}^{T} e_{t+1} \left(T^{-1} \sum_{j=1}^{t} e_j \right) / s \\ &\Longrightarrow -\sigma_e^2 \int_{\xi}^{1} s^{-1} W(s) \, \mathrm{d}W(s). \\ &\Longrightarrow -\sigma_e^2 \int_{0}^{1} s^{-1} W(s) \, \mathrm{d}W(s). \end{aligned}$$

□

Lemma 6. Under Assumptions 1b and 2c, $A_2 = \sum_{t=R}^{T} e_{t+1}\left[x_{2,t}'\left(\hat{\beta}_{2,t} - \beta_{2,t}\right)\right]$ is $O_p\left(1-\xi\right) = O_p\left(\frac{P}{T}\right)$ under $\mathbb{H}_0$.

Proof: Rewrite

$$\begin{aligned}
A_2 &= \sum_{t=R}^{T} e_{t+1}x_{2,t}'\left(\hat{\beta}_{2,t} - \beta_2\right) \\
&= \sum_{t=R}^{T} e_{t+1}x_{2,t}'\left(\sum_{j=1}^{t-1} x_{2,j}x_{2,j}'\right)^{-1} \times \left(\sum_{j=1}^{t-1} x_{2,j}e_{j+1}\right) \\
&= \sum_{t=R}^{T} e_{t+1}x_{2,t}'G_T^{-1}\left[G_T^{-1}\sum_{j=1}^{t-1} x_{2,j}x_{2,j}'G_T^{-1}\right]^{-1} \times \left[G_T^{-1}\sum_{j=1}^{t-1} x_{2,j}e_{j+1}\right],
\end{aligned}$$

where $G_T = \mathrm{diag}\big(T^{0.5}, T\big)$ as before, and for the two bracketed terms in the last line, we have

$$\begin{aligned}
G_T^{-1}\left(\sum_{j=1}^{t-1} x_{2,j}x_{2,j}'\right)G_T^{-1} &\Rightarrow \begin{pmatrix} s & \int_0^s J_x^c(r)\,\mathrm{d}r \\ \int_0^s J_x^c(r)\,\mathrm{d}r & \int_0^s \left(J_x^c(r)\right)^2\mathrm{d}r \end{pmatrix} \\
G_T^{-1}\sum_{j=1}^{t-1} x_{2,j}e_{j+1} &\Rightarrow \begin{pmatrix} \int_0^s 1\mathrm{d}W(r) \\ \int_0^s J_x^c(r)\,\mathrm{d}W(r) \end{pmatrix}
\end{aligned}$$

where $J_x^c(r)$ is an Ornstein-Uhlenbeck process, and $W(r)$ is a Wiener process. Hence

$$\begin{aligned}
A_2 &= \sum_{t=R}^{T} e_{t+1}x_{2,t}'G_T^{-1}\left[G_T^{-1}\sum_{j=1}^{t-1} x_{2,j}x_{2,j}'G_T^{-1}\right] \\
&\quad \times \left[G_T^{-1}\sum_{j=1}^{t-1} x_{2,j}e_{j+1}\right] \\
&\Rightarrow \int_\xi^1 \begin{pmatrix} 1 & J_x^c(s) \end{pmatrix} \begin{pmatrix} s & \int_0^s J_x^c(r)\,\mathrm{d}r \\ \int_0^s J_x^c(r)\,\mathrm{d}r & \int_0^s \left(J_x^c(r)\right)^2\mathrm{d}r \end{pmatrix}^{-1} \\
&\quad \times \begin{pmatrix} \int_0^s 1\mathrm{d}W(r) \\ \int_0^s J_x^c(r)\,\mathrm{d}W(r) \end{pmatrix} \mathrm{d}W(s)
\end{aligned}$$

Therefore $A_2 = \sum_{t=R}^{T} e_{t+1}x_{2,t}'\left(\hat{\beta}_{2,t} - \beta_{2,t}\right) = O_p\left(1-\xi\right)$. □

Remark 4. Note that the A_1 weakly converges to normal distribution, A_2 converges to another distribution, and $A_1 + A_2$ can not be written as did in Lemma A10 by Clark and McCracken (2001), that is, the t-statistic of the encompassing test is not standard normal.

Lemma 7. Under Assumptions 1b and 2c, $A_3 + A_4$ is $o_p(1)$ under $\mathbb{H}_0$.

Proof: Let $E_T = \operatorname{diag}\left(T^0, T^{0.5}\right)$, $F_T = \operatorname{diag}\left(T^1, T^{1.5}\right)$, $G_T = \operatorname{diag}\left(T^{0.5}, T^1\right)$, then for any 2×2 matrix K, we have $E_T F_T = G_T G_T$ and

$$E_T \times K \times F_T = G_T \times K \times G_T,$$

because E_T, F_T, G_T are diagonal. Therefore

$$\begin{aligned} A_3 + A_4 &= \sum\nolimits_{t=R}^{T} \left(\hat{\beta}_{1,t} - \beta_{1,t}\right) x_{1,t} x_{1,t}' \left(\hat{\beta}_{1,t} - \beta_{1,t}\right) \\ &\quad - \sum\nolimits_{t=R}^{T} \left(\hat{\beta}_{1,t} - \beta_{1,t}\right) x_{1,t} x_{2,t}' \left(\hat{\beta}_{2,t} - \beta_{2,t}\right) \\ &= \sum\nolimits_{t=R}^{T} H_1'(t) B_1(t) q_{1,t} B_1(t) H_1(t) - \sum\nolimits_{t=R}^{T} H_1'(t) B_1(t) x_{1,t} x_{2,t}' B_2(t) H_2(t), \end{aligned}$$

where the second line appears to be the same as the second bracketed right-hand side term in (A7) of Lemma A10 in CM (2001), which shows that the above is $o_p(1)$ under Assumption 1a. However, under Assumption 1b, x_t has an AR root local to unity. We show below that the local-to-unit root in x does not affect Lemma A10 of CM (2001). This is because terms involving x can be suitably normalized as follows

$$\begin{aligned} A_3 + A_4 &= \sum\nolimits_{t=R}^{T} H_1'(t) B_1(t) q_{1,t} B_1(t) H_1(t) \\ &\quad - \sum\nolimits_{t=R}^{T} H_1'(t) B_1(t) x_{1,t} \left(x_{2,t}' E_T^{-1}\right) \left[E_T \times t^{-1} B_2(t) \times F_T \times s\right] \\ &\quad \times \left[F_T^{-1} \times t H_2(t) / s\right] \\ &\equiv \sum\nolimits_{t=R}^{T} H_1'(t) B_1(t) q_{1,t} B_1(t) H_1(t) - \sum\nolimits_{t=R}^{T} H_1'(t) B_1(t) x_{1,t} \ddot{x}_{2,t}' \ddot{B}_2(t) \ddot{H}_2(t). \end{aligned}$$

where

$$\ddot{x}_{2,t}' \equiv x_{2,t}' E_T^{-1} \Rightarrow \begin{pmatrix} 1 & J_x^c(s) \end{pmatrix} = O_p(1),$$

$$
\begin{aligned}
\ddot{B}_2(t) &\equiv E_T \times t^{-1} B_2(t) \times F_T \times s \\
&= G_T \left[t^{-1} B_2(t)\right] G_T \times s \\
&= \left[G_T^{-1} \left[t^{-1} B_2(t)\right]^{-1} G_T^{-1}\right]_T^{-1} \times s \\
&= \left[G_T^{-1} \sum_{j=1}^{t} x_{2,j} x_{2,j}' G_T^{-1}\right]^{-1} \times s \\
&\Rightarrow \begin{pmatrix} s & \int_0^s J_x^c(r)\,\mathrm{d}r \\ \int_0^s J_x^c(r)\,\mathrm{d}r & \int_0^s \left(J_x^c(r)\right)^2 \mathrm{d}r \end{pmatrix}^{-1} \times s = O_p(1)
\end{aligned}
$$

for sufficiently small s, and

$$
\begin{aligned}
\ddot{H}_2(t) &\equiv F_T^{-1} \times R H_2(t)/s \\
&= F_T^{-1} \sum_{j=1}^{t-1} x_{2,j} e_{j+1}/s \\
&= \begin{pmatrix} T^{-1}/s \times \sum_{j=1}^{t-1} e_{j+1} \\ T^{-1.5}/s \times \sum_{j=1}^{t-1} x_j e_{j+1} \end{pmatrix} \\
&= \begin{pmatrix} T^{-0.5}/s \times T^{-0.5} \sum_{j=1}^{t-1} e_{j+1} \\ T^{-0.5}/s \times T^{-1} \sum_{j=1}^{t-1} x_j e_{j+1} \end{pmatrix} \\
&\Rightarrow \begin{pmatrix} T^{-0.5}/s \times \int_0^s 1 \mathrm{d}W(r) \\ T^{-0.5}/s \times \int_0^s J_x^c(r)\,\mathrm{d}W(r) \end{pmatrix} \\
&= O\left(T^{-0.5}\right).
\end{aligned}
$$

when s is sufficiently small. Therefore, $\ddot{x}_{2,t}', \ddot{B}_2(t), \ddot{H}_2(t)$ have the same orders of magnitude as $x_{2,t}', B_2(t), H_2(t)$ in stationary case of Lemma A10 in CM (2001). Therefore $A_3 + A_4$ is $o(1)$ not only under Assumption 1a but also under Assumption 1b. □

Remark 5. Based on Lemmas 5-7 under Assumptions 1b and 2c, $\sum_{t=R}^{T} \hat{e}_{t+1}^{(1)} \left(\hat{e}_{t+1}^{(1)} - \hat{e}_{t+1}^{(2)}\right) = A_1 + A_2 + o_p(1)$ as summarized in Proposition 7 for the recursive scheme. However, here, $A_1 + A_2$ can not be written as in Lemma A10 of CM (2001), and therefore the t statistic of the encompassing test is *not* standard normal.

6 Asymptotic Distribution of ENC When There is a Contemporaneous Correlation

Recall the big model

$$\begin{aligned} y_{t+1} &= c_2 + bx_t + e_{t+1}^{(2)} \\ x_{t+1} &= \phi x_t + v_{t+1} \end{aligned}$$

where $\rho(e_t, v_t) \neq 0$. Note that both $\{v_t\}$ and $\{e_t\}$ are i.i.d, we have $Cov(v_m, e_n) = 0$ when $m \neq n$. Therefore

$$Cov(x_t, e_{t+1}) = Cov((1 - \phi L) v_t, e_{e+1}) = 0,$$

where L is the lag operator, Therefore, $\left(\hat{c}_{2,t}, \hat{b}_t\right)$ is unbiased and consistent. In fact, the coefficients are solved using the following restricted VAR model

$$\begin{pmatrix} y_{t+1} \\ x_{t+1} \end{pmatrix} = \begin{pmatrix} c_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & b \\ 0 & \phi \end{pmatrix} \begin{pmatrix} y_t \\ x_t \end{pmatrix} + \begin{pmatrix} e_{t+1}^{(2)} \\ v_{t+1} \end{pmatrix}$$

Since the matrix $\begin{pmatrix} 0 & b \\ 0 & \phi \end{pmatrix}$ is restricted matrix (only two parameters to be estimated), we are unable to estimate the coefficients by running regression for each line. In Monte-Carlo simulation, we set $Corr(v_t, e_t) = 0.5$.

7 Monte Carlo

We compare the two statistics by changing $\{x_t\}$ from stationary process to Ornstein-Uhlenbeck process and use the data generating process in Model 1 and Model 2 as follows: the additional variable x_t in Model 2 has the AR process: $x_t = \phi x_{t-1} + z_t$, where z_j is i.i.d, following $N(0,1)$, and $\mathbb{E}(z_t | x_{t-1}) = 0$, The error term $e_{t+1}^{(2)} \sim N\left(0, \sigma_e^2\right)$. We set $c_2 = 1$ (we also set $c_2 = 0$ for martingale difference series $\{y_t\}$ and report the result.) $b \in \{0.0, 0.1, 1.0\}$, $\phi \in \{0, 0.5, 0.9, 0.95, 0.99, 1\}$, and $\sigma_e \in \{0.1, 1.0\}$. Model 1 is estimated by regressing $\{y_j\}_{j=t-R+1}^{t}$ on constant term to obtain $\hat{c}_{1,t}$, where $t = R, \ldots, T$. Model 2 is estimated by regressing $\{y_j\}_{j=t-R+1}^{t}$ on $\{1, x_{j-1}\}_{j=t-R+1}^{t}$ to obtain $\left(\hat{c}_{2,t}\ \hat{b}_t\right)$. The forecast errors from the two models are $\hat{e}_{t+1}^{(1)} = y_{t+1} - \hat{c}_{1,t}$ and $\hat{e}_{t+1}^{(2)} = y_{t+1} - \hat{c}_{2,t} - \hat{b}_t x_t$ over the forecast evaluation period at $t = R, \ldots, T$. The number of observations for the rolling windows for estimation are chosen from $R \in \{60, 120, 240\}$. Let $P = T - R + 1 \in \{48, 240, 1200\}$. From these, we compute the two statistics DM_P and ENC_P. All above tests are repeated 2000 times to find out the Monte Carlo distributions of DM_P and ENC_P, and to compute their size and power. Also we consider the test statistics when $\rho \equiv Corr\left(v_t, e_{t+1}\right) = -0.95$, $R_T = 0.8$.

The Table 1 shows the size of test with additional covariate from stationary process to local to Ornstein-Uhlenbeck process using rolling scheme when $c_1 \neq 0$. We see that DM statistics are undersized for large P/R ratio. ENC test is robust, having correct size for all ϕ ranging from 0 to 0.99. Table 2 and 3 show the power of test using rolling scheme when $c_1 \neq 0$. We see that as ϕ increases, the powers approach to 1 dramatically since higher ϕ implies higher singal-to-noise ratio. Tables 4-6 are similar to table 1-3, showing the statistics of DM and ENC test when there is no constant term in Model 1, we see that ENC test also has a good performance since the test is asymptotically standard normal under null hypothesis. Tables 7-9 show the size (power) of test with additional covariate from stationary process to local to Ornstein-Uhlenbeck process using rolling scheme when $c_1 \neq 0$ and there is existing a contemporaneous correlation that $\rho\left(v_t, e_t\right) = 0.5$. We see that ENC also has good size, since $\rho\left(v_t, e_t\right) \neq 0$ does not necessarily mean that $\rho\left(v_t, e_{t+1}\right) \neq 0$,therefore the encompassing test is regardless of ϕ. Tables 10-12 show the size (power) of test using recursive scheme. Figures 1-12 show the Monte Carlo distributions of DM_P and ENC_P using rolling scheme when $c_1 \neq 0$. Under $\mathbb{H}_0$, we can see that for size of test, ENC has the correct size regardless of ϕ. Figures 13-24 show the Monte Carlo distributions of DM_P and ENC_P using rolling scheme when $c_1 = 0$. Figures 25-36 show the Monte Carlo distributions of DM_P and ENC_P using rolling scheme when $c_1 \neq 0$ with contemporaneous correlation $\rho\left(e_t, v_t\right) = 0.5$, we see that ENC also have a good size. Figures 25-36 show the Monte Carlo distributions using recursive scheme.

Tables 1-12 About Here

Figures 1-48 About Here

8 Application

We apply the two statistics (DM, ENC) to the Goyal and Welch study (2008) and construct nested models to test if a predictor such as dividend-yield ratio (DY), dividend-price ratio (DP), long term rate of yield (LTY) and inflation (INFL) Granger-causes the equity premium in the conditional mean. The dividend-yield ratio at time t is defined as the most recent dividend at t divided by stock price at time t, the dividend-price ratio at time t is defined as the most recent dividend at $t-1$ divided by stock price at time t. The explanation of other variables are available from the homepage of A. Goyal. The two nested models are as shown in (12) and (13) below

$$\text{Model 1} \quad : \quad y_{t+1} = c_1 + e_{t+1}^{(1)}, \tag{12}$$

$$\text{Model 2} \quad : \quad y_{t+1} = c_2 + bx_t + e_{t+1}^{(2)}, \tag{13}$$

where y_{t+1} is the equity premium and x_t is the covariate. We use monthly data ranging from 1926 to 2011, containing 1032 observation for all four models. In Model 1, we only have a constant term, therefore at time t, we predict the future equity premium by solely using the historical average of previous R observations of the equity premium from time $t-R+1$ to t. In Model 2, we use the 1-lag covariate to forecast the equity premium in the next month. See Goyal and Welch (2008) for more on data descriptions. We intend to check if two nested models have the same predictive accuracy. We use the rolling window scheme of the window size R starting from the 15% of the total observations is $T = 1032$ to the 85% of T. The in-sample observation R ranges from $R = 155$ ($R/P = 155/877$) to $R = 877$ ($R/P = 877/155$). The red line represents DM_P under different allocation of R and P. The blue line represents ENC_P.

We find: (i) The ENC statistics always have higher statistics than DM test. (ii) DP and DY figures show that ENC test is significant with small R/P ratio. Intuitively, for R/P, we are unable to account for the y_{t+1} by solely using the previous y up to time t since there is not sufficient information available, therefore we need to exploit the property of additional variable x. In this way, x has predictive power for ENC test; however when R/P is large, we can predict y_{t+1} using previous information of y up to time t, which weakens the predictive power of x. (iii) LTY figure shows that long-term yield has predictive power for equity premium for all R ranging from 150 to 877 using ENC test, which can not directly been observed from DM statistic. (iv) INFL figure shows that for ENC test, the inflation rate has predictive power when the number of in-sample observations is below 600.

Figure 49 About Here

9 Conclusions

This paper extends the work of CM (2001), CW (2006, 2007) from nested mean model with weak stationary predictor to nested mean model with a highly persistent predictor. CM (2001) and CW (2006, 2007) found that the DM statistic tends to be negative under the null hypothesis of the equal predictive ability because of parameter estimation error. We find that the DM is even more severely undersized when the predictor is highly persistent with the AR root closer to unity. We find that the ENC test is robust as it remains the correct size under the null hypothesis of the equal predictive ability, also it has high power under alternative. We show that the highly persistent predictor following the Ornstein-Uhlenbeck process implies the convergent rate of the estimator from Model 2 faster than that from a model with stationary predictor and the ENC test can shown to be asymptotically standard normal when the ratio of out-of-sample to in-sample observation is infinite. By using Monte-Carlo simulation, we see that ENC test is robust and has the correct size under null hypothesis whereas DM is severely under-sized. An application to the predictive regression of the equity premium reveals strong predictive ability of several persistent predictors.

References

Bobkoski, M.J. (1983), "Hypothesis Testing in Nonstationary Time Series", Unpublished PhD thesis, Dept. of Statistics, University of Wisconsin, Madison.

Breitung, J. and M. Demetrescub (2015), "Instrumental variable and variable addition based inference in predictive regressions", *Journal of Econometrics* 187(1): 358-375.

Cai, Z. and Y. Wang (2014), "Testing predictive regression models with nonstationary regressors", *Journal of Econometrics* 178(1): 4-14. Corrigendum (2014) *Journal of Econometrics* 181: page 194.

Cavanagh, C. (1985), "Roots Local to Unity, Unpublished Manuscript", Harvard University.

Chan, N. H. and C. Z. Wei (1987), "Asymptotic Inference for Nearly Nonstationary AR(1) Processes", *Annals of Statistics* 15(3), 1050-1063.

Clark, T. E. and M. McCracken (2001). "Tests of Equal Forecast Accuracy and Encompassing for Nested Models." *Journal of Econometrics* 105(1): 85-110.

Clark, T. E. and K. D. West (2006). "Using Out-of-sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis." *Journal of Econometrics* 135(1-2): 155-186.

Clark, T. E. and K. D. West (2007). "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models." *Journal of Econometrics* 138(1): 291-311.

Diebold, F.X. (2015), "Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold–Mariano Tests.", *Journal of Business and Economic Statistics* 33(1): 1-8.

Diebold, F. X. and R. S. Mariano (1995). "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics* 13(3): 253-263.

Giacomini, R. and H. White (2006). "Tests of Conditional Predictive Ability." *Econometrica* 74(6): 1545-1578.

Giraitis, L. and P. C. B. Phillips (2006), "Uniform Limit Theory for Stationary Autoregression", *Journal of Time Series Analysis* 27(1), 51-60.

Goyal, A. and I. Welch (2008). "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies* 21: 1455-1508.

Harvey, D. S., S. J. Leybourne and P. Newbold (1998). "Tests for Forecast Encompassing." *Journal of Business and Economic Statistics* 16(2): 254-259.

Jansson, M. and M. Moreira (2006), "Optimal Inference in Regression Models with Nearly Integrated Regressors", *Econometrica* 74(3): 681-714.

Mikusheva, A. (2007), "Uniform inference in autoregressive models," *Econometrica* 75(5): 1411-1452.

Mikusheva, A. (2015), "Second Order Expansion of the t-statistic in AR(1) Models," *Econometric Theory* 31: 426-448.

Park, J. Y. (2003): "Weak Unit Roots", Unpublished Manuscript.

Phillips, P. C. B. (1987a), "Time Series Regression with a Unit Root", *Econometrica* 55(2): 277-302.

Phillips, P. C. B. (1987b), "Towards a Unified Asymptotic Theory for Autoregression", *Biometrika* 74(3): 535-547.

Phillips, P. C. B. (2014), "On Confidence Intervals for Autoregressive Roots and Predictive Regression", *Econometrica* 82(3): 1177-1195.

Phillips, P. C. B. and J. H. Lee (2013), "Predictive Regression Under Various Degrees of Persistence and Robust Long-Horizon Regression", *Journal of Econometrics* 177(2): 250-264.

Phillips, P. C. B. and J. H. Lee (2014), "Robust Econometric Inference with Mixed Integrated and Mildly Explosive Regressors", Working Paper.

Phillips, P. C. B. and T. Magdalinos (2007), "Limit Theory for Moderate Deviations from a Unit Root", *Journal of Econometrics* 137(1): 115-130.

Phillips, P. C. B. and T. Magdalinos (2009), "Econometric Inference in the Vicinity of Unity", Working Paper.

Stambaugh, R. F. (1999), "Predictive Regressions", *Journal of Financial Economics* 54(3): 375-421.

Stock, J. H. (1991). "Confidence Intervals for the Largest Autoregressive Root in US Macroeconomic Time Series." *Journal of Monetary Economics* 28: 435-459.

West, K. D. (1996). "Asymptotic Inference about Predictive Ability." *Econometrica* 64(5): 1067-1084.

Table 1: Rejection frequency at 5% level, $b = 0$ (With intercept Model 1l)

		$P = 48$		$P = 240$		$P = 1200$	
$Repeat = 2000$		DM_P	ENC_P	DM_P	ENC_P	DM_P	ENC_P
$\rho = 0,\ \sigma_e = 0.1$	$R = 60$	0.010	0.039	0.000	0.037	0.000	0.043
	$R = 120$	0.009	0.030	0.001	0.029	0.000	0.042
	$R = 240$	0.025	0.040	0.006	0.034	0.000	0.032
$\rho = 0,\ \sigma_e = 1$	$R = 60$	0.005	0.027	0.000	0.026	0.000	0.030
	$R = 120$	0.016	0.037	0.000	0.029	0.000	0.038
	$R = 240$	0.029	0.046	0.006	0.025	0.000	0.031
$\rho = 0.1,\ \sigma_e = 0.1$	$R = 60$	0.012	0.049	0.000	0.037	0.000	0.040
	$R = 120$	0.018	0.040	0.000	0.029	0.000	0.031
	$R = 240$	0.018	0.033	0.006	0.030	0.000	0.029
$\rho = 0.1,\ \sigma_e = 1$	$R = 60$	0.008	0.036	0.000	0.037	0.000	0.040
	$R = 120$	0.018	0.038	0.002	0.034	0.000	0.034
	$R = 240$	0.025	0.037	0.003	0.029	0.000	0.031
$\rho = 0.5,\ \sigma_e = 0.1$	$R = 60$	0.005	0.033	0.000	0.028	0.000	0.038
	$R = 120$	0.015	0.036	0.001	0.026	0.000	0.036
	$R = 240$	0.018	0.032	0.004	0.028	0.000	0.032
$\rho = 0.5,\ \sigma_e = 1$	$R = 60$	0.010	0.038	0.001	0.036	0.000	0.044
	$R = 120$	0.019	0.036	0.000	0.028	0.000	0.039
	$R = 240$	0.023	0.039	0.004	0.029	0.000	0.035
$\rho = 0.9,\ \sigma_e = 0.1$	$R = 60$	0.007	0.035	0.000	0.033	0.000	0.043
	$R = 120$	0.011	0.034	0.001	0.026	0.000	0.044
	$R = 240$	0.023	0.044	0.004	0.024	0.000	0.033
$\rho = 0.9,\ \sigma_e = 1$	$R = 60$	0.007	0.035	0.000	0.034	0.000	0.048
	$R = 120$	0.016	0.035	0.000	0.027	0.000	0.037
	$R = 240$	0.023	0.040	0.003	0.032	0.000	0.036
$\rho = 0.95,\ \sigma_e = 0.1$	$R = 60$	0.002	0.024	0.000	0.028	0.000	0.047
	$R = 120$	0.012	0.036	0.001	0.036	0.000	0.039
	$R = 240$	0.023	0.038	0.005	0.032	0.000	0.036
$\rho = 0.95,\ \sigma_e = 1$	$R = 60$	0.005	0.028	0.000	0.037	0.000	0.040
	$R = 120$	0.013	0.033	0.001	0.027	0.000	0.047
	$R = 240$	0.017	0.031	0.003	0.017	0.000	0.034
$\rho = 0.99,\ \sigma_e = 0.1$	$R = 60$	0.008	0.034	0.000	0.031	0.000	0.037
	$R = 120$	0.012	0.036	0.001	0.034	0.000	0.043
	$R = 240$	0.026	0.049	0.001	0.019	0.000	0.040
$\rho = 0.99,\ \sigma_e = 1$	$R = 60$	0.003	0.033	0.000	0.032	0.000	0.045
	$R = 120$	0.010	0.032	0.000	0.031	0.000	0.040
	$R = 240$	0.021	0.036	0.003	0.030	0.000	0.035

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Rolling Scheme).

Table 2: Rejection frequency at 5% level, $b = 0.1$ (With intercept on Model 1l)

Repeat = 2000		$P = 48$		$P = 240$		$P = 1200$	
		DM_P	ENC_P	DM_P	ENC_P	DM_P	ENC_P
$\rho = 0$, $\sigma_e = 0.1$	$R = 60$	0.954	1.000	1.000	1.000	1.000	1.000
	$R = 120$	0.948	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.943	1.000	1.000	1.000	1.000	1.000
$\rho = 0$, $\sigma_e = 1$	$R = 60$	0.023	0.107	0.004	0.233	0.001	0.601
	$R = 120$	0.049	0.124	0.015	0.247	0.021	0.728
	$R = 240$	0.070	0.142	0.060	0.325	0.129	0.823
$\rho = 0.1$, $\sigma_e = 0.1$	$R = 60$	0.949	1.000	1.000	1.000	1.000	1.000
	$R = 120$	0.943	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.945	1.000	1.000	1.000	1.000	1.000
$\rho = 0.1$, $\sigma_e = 1$	$R = 60$	0.020	0.097	0.006	0.214	0.000	0.614
	$R = 120$	0.040	0.108	0.016	0.255	0.032	0.731
	$R = 240$	0.069	0.149	0.055	0.331	0.110	0.829
$\rho = 0.5$, $\sigma_e = 0.1$	$R = 60$	0.983	1.000	1.000	1.000	1.000	1.000
	$R = 120$	0.977	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.975	1.000	1.000	1.000	1.000	1.000
$\rho = 0.5$, $\sigma_e = 1$	$R = 60$	0.032	0.122	0.006	0.284	0.002	0.751
	$R = 120$	0.057	0.159	0.031	0.347	0.064	0.848
	$R = 240$	0.081	0.164	0.079	0.404	0.238	0.923
$\rho = 0.9$, $\sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	0.999	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.9$, $\sigma_e = 1$	$R = 60$	0.078	0.311	0.128	0.771	0.511	1.000
	$R = 120$	0.133	0.390	0.301	0.865	0.936	1.000
	$R = 240$	0.170	0.445	0.435	0.933	0.993	1.000
$\rho = 0.95$, $\sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.95$, $\sigma_e = 1$	$R = 60$	0.110	0.441	0.317	0.927	0.926	1.000
	$R = 120$	0.220	0.571	0.619	0.976	1.000	1.000
	$R = 240$	0.255	0.586	0.709	0.990	1.000	1.000
$\rho = 0.99$, $\sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.99$, $\sigma_e = 1$	$R = 60$	0.252	0.623	0.676	0.986	1.000	1.000
	$R = 120$	0.434	0.770	0.928	1.000	1.000	1.000
	$R = 240$	0.519	0.824	0.970	1.000	1.000	1.000

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Rolling Scheme).

Table 3: Rejection frequency at 5% level, $b = 1$ (With intercept on Model 1)

		$P = 48$		$P = 240$		$P = 1200$	
$Repeat = 2000$		DM_P	ENC_P	DM_P	ENC_P	DM_P	ENC_P
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.964	1.000	1.000	1.000	1.000	1.000
$\rho = 0$, $\sigma_e = 1$	$R = 120$	0.950	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.953	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.1$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.953	1.000	1.000	1.000	1.000	1.000
$\rho = 0.1$, $\sigma_e = 1$	$R = 120$	0.956	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.939	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.5$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.979	1.000	1.000	1.000	1.000	1.000
$\rho = 0.5$, $\sigma_e = 1$	$R = 120$	0.980	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.970	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.9$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.9$, $\sigma_e = 1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.95$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.95$, $\sigma_e = 1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.99$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.99$, $\sigma_e = 1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Rolling Scheme).

Table 4: Rejection frequency at 5% level, $b = 0$ (Without intercept on Model 1)

		$P = 48$		$P = 240$		$P = 1200$	
$Repeat = 2000$		DM_P	ENC_P	DM_P	ENC_P	DM_P	ENC_P
	$R = 60$	0.006	0.042	0.000	0.045	0.000	0.047
$\rho = 0$, $\sigma_e = 0.1$	$R = 120$	0.016	0.040	0.000	0.033	0.000	0.047
	$R = 240$	0.019	0.042	0.002	0.035	0.000	0.032
	$R = 60$	0.004	0.032	0.000	0.035	0.000	0.048
$\rho = 0$, $\sigma_e = 1$	$R = 120$	0.014	0.043	0.001	0.032	0.000	0.038
	$R = 240$	0.020	0.051	0.003	0.039	0.000	0.032
	$R = 60$	0.003	0.033	0.000	0.035	0.000	0.038
$\rho = 0.1$, $\sigma_e = 0.1$	$R = 120$	0.013	0.046	0.001	0.026	0.000	0.045
	$R = 240$	0.019	0.038	0.002	0.034	0.000	0.042
	$R = 60$	0.006	0.042	0.000	0.031	0.000	0.042
$\rho = 0.1$, $\sigma_e = 1$	$R = 120$	0.013	0.046	0.001	0.029	0.000	0.041
	$R = 240$	0.023	0.052	0.003	0.030	0.000	0.040
	$R = 60$	0.006	0.038	0.000	0.035	0.000	0.041
$\rho = 0.5$, $\sigma_e = 0.1$	$R = 120$	0.010	0.039	0.000	0.025	0.000	0.036
	$R = 240$	0.018	0.041	0.001	0.033	0.000	0.034
	$R = 60$	0.006	0.039	0.000	0.038	0.000	0.036
$\rho = 0.5$, $\sigma_e = 1$	$R = 120$	0.009	0.034	0.000	0.031	0.000	0.040
	$R = 240$	0.017	0.036	0.002	0.029	0.000	0.034
	$R = 60$	0.003	0.039	0.000	0.040	0.000	0.047
$\rho = 0.9$, $\sigma_e = 0.1$	$R = 120$	0.008	0.037	0.000	0.038	0.000	0.036
	$R = 240$	0.023	0.043	0.002	0.032	0.000	0.038
	$R = 60$	0.003	0.032	0.000	0.044	0.000	0.045
$\rho = 0.9$, $\sigma_e = 1$	$R = 120$	0.008	0.040	0.000	0.029	0.000	0.036
	$R = 240$	0.018	0.040	0.002	0.028	0.000	0.035
	$R = 60$	0.003	0.031	0.000	0.036	0.000	0.052
$\rho = 0.95$, $\sigma_e = 0.1$	$R = 120$	0.005	0.030	0.000	0.034	0.000	0.042
	$R = 240$	0.018	0.038	0.002	0.031	0.000	0.037
	$R = 60$	0.003	0.038	0.000	0.040	0.000	0.055
$\rho = 0.95$, $\sigma_e = 1$	$R = 120$	0.012	0.042	0.000	0.028	0.000	0.037
	$R = 240$	0.011	0.032	0.002	0.032	0.000	0.031
	$R = 60$	0.003	0.028	0.000	0.035	0.000	0.049
$\rho = 0.99$, $\sigma_e = 0.1$	$R = 120$	0.008	0.039	0.000	0.035	0.000	0.040
	$R = 240$	0.014	0.035	0.002	0.031	0.000	0.036
	$R = 60$	0.002	0.027	0.000	0.036	0.000	0.045
$\rho = 0.99$, $\sigma_e = 1$	$R = 120$	0.005	0.030	0.000	0.028	0.000	0.045
	$R = 240$	0.015	0.037	0.002	0.023	0.000	0.030

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Rolling Scheme).

Table 5: Rejection frequency at 5% level, $b = 0.1$. (Without intercept on Model 1)

		$P = 48$		$P = 240$		$P = 1200$	
$Repeat = 2000$		DM_P	ENC_P	DM_P	ENC_P	DM_P	ENC_P
	$R = 60$	0.937	1.000	1.000	1.000	1.000	1.000
$\rho = 0$, $\sigma_e = 0.1$	$R = 120$	0.948	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.943	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.013	0.078	0.001	0.157	0.000	0.463
$\rho = 0$, $\sigma_e = 1$	$R = 120$	0.037	0.103	0.006	0.225	0.001	0.595
	$R = 240$	0.050	0.131	0.026	0.268	0.031	0.758
	$R = 60$	0.932	1.000	1.000	1.000	1.000	1.000
$\rho = 0.1$, $\sigma_e = 0.1$	$R = 120$	0.940	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.935	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.014	0.073	0.002	0.182	0.000	0.484
$\rho = 0.1$, $\sigma_e = 1$	$R = 120$	0.031	0.112	0.007	0.214	0.001	0.624
	$R = 240$	0.051	0.124	0.035	0.263	0.027	0.731
	$R = 60$	0.973	1.000	1.000	1.000	1.000	1.000
$\rho = 0.5$, $\sigma_e = 0.1$	$R = 120$	0.974	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.973	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.010	0.102	0.001	0.230	0.000	0.616
$\rho = 0.5$, $\sigma_e = 1$	$R = 120$	0.038	0.129	0.012	0.296	0.007	0.753
	$R = 240$	0.062	0.150	0.051	0.366	0.075	0.873
	$R = 60$	0.998	1.000	1.000	1.000	1.000	1.000
$\rho = 0.9$, $\sigma_e = 0.1$	$R = 120$	0.999	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.998	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.052	0.298	0.050	0.753	0.119	1.000
$\rho = 0.9$, $\sigma_e = 1$	$R = 120$	0.099	0.356	0.191	0.838	0.726	1.000
	$R = 240$	0.144	0.401	0.369	0.902	0.967	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.95$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.114	0.468	0.264	0.941	0.818	1.000
$\rho = 0.95$, $\sigma_e = 1$	$R = 120$	0.194	0.552	0.521	0.972	0.995	1.000
	$R = 240$	0.228	0.577	0.678	0.989	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.99$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.473	0.795	0.875	0.998	1.000	1.000
$\rho = 0.99$, $\sigma_e = 1$	$R = 120$	0.544	0.831	0.953	0.999	1.000	1.000
	$R = 240$	0.547	0.827	0.968	1.000	1.000	1.000

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Rolling Scheme).

Table 6: Rejection frequency at 5% level, $b = 1$. (Without intercept on Model 1)

$Repeat = 2000$		$P = 48$ DM_P	$P = 48$ ENC_P	$P = 240$ DM_P	$P = 240$ ENC_P	$P = 1200$ DM_P	$P = 1200$ ENC_P
$\rho = 0,\ \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0,\ \sigma_e = 1$	$R = 60$	0.939	1.000	1.000	1.000	1.000	1.000
	$R = 120$	0.940	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.939	1.000	1.000	1.000	1.000	1.000
$\rho = 0.1,\ \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.1,\ \sigma_e = 1$	$R = 60$	0.934	1.000	1.000	1.000	1.000	1.000
	$R = 120$	0.939	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.940	1.000	1.000	1.000	1.000	1.000
$\rho = 0.5,\ \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.5,\ \sigma_e = 1$	$R = 60$	0.975	1.000	1.000	1.000	1.000	1.000
	$R = 120$	0.976	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.970	1.000	1.000	1.000	1.000	1.000
$\rho = 0.9,\ \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.9,\ \sigma_e = 1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	0.999	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.999	1.000	1.000	1.000	1.000	1.000
$\rho = 0.95,\ \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.95,\ \sigma_e = 1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.99,\ \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\rho = 0.99,\ \sigma_e = 1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Rolling Scheme).

Table 7: Rejection frequency at 5% level, $b = 0$, $\rho(e_t, v_t) = 0.5$ (With intercept on Model 1)

		$P = 48$		$P = 240$		$P = 1200$	
$Repeat = 2000$		DM_P	ENC_P	DM_P	ENC_P	DM_P	ENC_P
	$R = 60$	0.005	0.039	0.000	0.024	0.000	0.050
$\phi = 0$, $\sigma_e = 0.1$	$R = 120$	0.018	0.037	0.002	0.027	0.000	0.034
	$R = 240$	0.021	0.038	0.003	0.031	0.000	0.031
	$R = 60$	0.008	0.037	0.000	0.032	0.000	0.043
$\phi = 0$, $\sigma_e = 1$	$R = 120$	0.013	0.032	0.001	0.030	0.000	0.038
	$R = 240$	0.027	0.043	0.005	0.033	0.000	0.036
	$R = 60$	0.010	0.034	0.000	0.029	0.000	0.037
$\phi = 0.1$, $\sigma_e = 0.1$	$R = 120$	0.012	0.034	0.001	0.029	0.000	0.039
	$R = 240$	0.018	0.036	0.005	0.030	0.000	0.032
	$R = 60$	0.012	0.033	0.000	0.028	0.000	0.047
$\phi = 0.1$, $\sigma_e = 1$	$R = 120$	0.013	0.032	0.002	0.031	0.000	0.032
	$R = 240$	0.026	0.045	0.006	0.028	0.000	0.040
	$R = 60$	0.011	0.036	0.001	0.023	0.000	0.042
$\phi = 0.5$, $\sigma_e = 0.1$	$R = 120$	0.014	0.031	0.000	0.029	0.000	0.031
	$R = 240$	0.019	0.034	0.004	0.020	0.000	0.031
	$R = 60$	0.008	0.029	0.000	0.033	0.000	0.039
$\phi = 0.5$, $\sigma_e = 1$	$R = 120$	0.017	0.036	0.001	0.029	0.000	0.048
	$R = 240$	0.019	0.036	0.002	0.023	0.001	0.037
	$R = 60$	0.006	0.039	0.001	0.041	0.000	0.047
$\phi = 0.9$, $\sigma_e = 0.1$	$R = 120$	0.013	0.034	0.002	0.025	0.000	0.045
	$R = 240$	0.024	0.039	0.003	0.024	0.000	0.030
	$R = 60$	0.007	0.030	0.000	0.033	0.000	0.040
$\phi = 0.9$, $\sigma_e = 1$	$R = 120$	0.017	0.037	0.000	0.027	0.000	0.046
	$R = 240$	0.024	0.040	0.006	0.031	0.000	0.029
	$R = 60$	0.008	0.042	0.000	0.037	0.000	0.042
$\phi = 0.95$, $\sigma_e = 0.1$	$R = 120$	0.014	0.032	0.001	0.040	0.000	0.045
	$R = 240$	0.031	0.046	0.003	0.030	0.000	0.036
	$R = 60$	0.010	0.041	0.000	0.037	0.000	0.067
$\phi = 0.95$, $\sigma_e = 1$	$R = 120$	0.020	0.043	0.001	0.036	0.000	0.038
	$R = 240$	0.032	0.049	0.003	0.026	0.000	0.035
	$R = 60$	0.006	0.043	0.001	0.045	0.000	0.111
$\phi = 0.99$, $\sigma_e = 0.1$	$R = 120$	0.023	0.051	0.001	0.036	0.000	0.064
	$R = 240$	0.030	0.051	0.003	0.027	0.000	0.040
	$R = 60$	0.004	0.037	0.001	0.052	0.000	0.098
$\phi = 0.99$, $\sigma_e = 1$	$R = 120$	0.022	0.056	0.001	0.039	0.000	0.061
	$R = 240$	0.030	0.052	0.005	0.037	0.000	0.041

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Rolling Scheme with Contemporaneous Correlation).

Table 8: Rejection frequency at 5% level, $b = 0.1$, $\rho(e_t, v_t) = 0.5$ (With intercept on Model 1)

$Repeat = 2000$		$P = 48$		$P = 240$		$P = 1200$	
		DM_P	ENC_P	DM_P	ENC_P	DM_P	ENC_P
	$R = 60$	0.966	1.000	1.000	1.000	1.000	1.000
$\phi = 0$, $\sigma_e = 0.1$	$R = 120$	0.954	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.948	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.021	0.096	0.007	0.214	0.001	0.566
$\phi = 0$, $\sigma_e = 1$	$R = 120$	0.037	0.120	0.019	0.245	0.027	0.710
	$R = 240$	0.066	0.135	0.060	0.315	0.125	0.818
	$R = 60$	0.959	1.000	1.000	1.000	1.000	1.000
$\phi = 0.1$, $\sigma_e = 0.1$	$R = 120$	0.953	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.948	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.022	0.091	0.005	0.198	0.004	0.576
$\phi = 0.1$, $\sigma_e = 1$	$R = 120$	0.045	0.110	0.023	0.254	0.028	0.733
	$R = 240$	0.063	0.140	0.065	0.336	0.136	0.846
	$R = 60$	0.974	1.000	1.000	1.000	1.000	1.000
$\phi = 0.5$, $\sigma_e = 0.1$	$R = 120$	0.972	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.973	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.027	0.116	0.013	0.258	0.007	0.688
$\phi = 0.5$, $\sigma_e = 1$	$R = 120$	0.049	0.128	0.033	0.327	0.068	0.841
	$R = 240$	0.068	0.143	0.080	0.404	0.241	0.912
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.9$, $\sigma_e = 0.1$	$R = 120$	0.996	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.996	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.065	0.264	0.141	0.679	0.441	0.994
$\phi = 0.9$, $\sigma_e = 1$	$R = 120$	0.136	0.350	0.315	0.811	0.909	1.000
	$R = 240$	0.163	0.370	0.450	0.874	0.983	1.000
	$R = 60$	0.999	1.000	1.000	1.000	1.000	1.000
$\phi = 0.95$, $\sigma_e = 0.1$	$R = 120$	0.998	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.999	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.099	0.370	0.283	0.822	0.801	1.000
$\phi = 0.95$, $\sigma_e = 1$	$R = 120$	0.199	0.475	0.580	0.935	0.998	1.000
	$R = 240$	0.253	0.541	0.690	0.965	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.99$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.168	0.504	0.502	0.933	0.977	1.000
$\phi = 0.99$, $\sigma_e = 1$	$R = 120$	0.387	0.679	0.871	0.985	1.000	1.000
	$R = 240$	0.494	0.760	0.947	0.997	1.000	1.000

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Rolling Scheme with Contemporaneous Correlation).

Table 9: Rejection frequency at 5% level, $b = 1$, $\rho(e_t, v_t) = 0.5$ (With intercept on Model 1)

		$P = 48$		$P = 240$		$P = 1200$	
$Repeat = 2000$		DM_P	ENC_P	DM_P	ENC_P	DM_P	ENC_P
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.951	1.000	1.000	1.000	1.000	1.000
$\phi = 0$, $\sigma_e = 1$	$R = 120$	0.953	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.955	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.1$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.952	1.000	1.000	1.000	1.000	1.000
$\phi = 0.1$, $\sigma_e = 1$	$R = 120$	0.947	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.949	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.5$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.976	1.000	1.000	1.000	1.000	1.000
$\phi = 0.5$, $\sigma_e = 1$	$R = 120$	0.963	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.969	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.9$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.998	1.000	1.000	1.000	1.000	1.000
$\phi = 0.9$, $\sigma_e = 1$	$R = 120$	0.999	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.996	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.95$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.95$, $\sigma_e = 1$	$R = 120$	0.999	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.99$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.999	1.000	1.000	1.000	1.000	1.000
$\phi = 0.99$, $\sigma_e = 1$	$R = 120$	0.999	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Rolling Scheme with Contemporaneous Correlation).

Table 10: Rejection frequency at 5% level, $b = 0$ (With intercept on Model 1)

		$P = 48$		$P = 240$		$P = 1200$	
$Repeat = 2000$		DM_P	ENC_P	DM_P	ENC_P	DM_P	ENC_P
	$R = 60$	0.006	0.027	0.002	0.025	0.000	0.024
$\phi = 0$, $\sigma_e = 0.1$	$R = 120$	0.018	0.042	0.002	0.032	0.001	0.027
	$R = 240$	0.027	0.042	0.005	0.031	0.000	0.025
	$R = 60$	0.011	0.034	0.002	0.027	0.001	0.028
$\phi = 0$, $\sigma_e = 1$	$R = 120$	0.019	0.031	0.005	0.030	0.000	0.020
	$R = 240$	0.028	0.041	0.006	0.034	0.002	0.027
	$R = 60$	0.007	0.038	0.001	0.029	0.001	0.027
$\phi = 0.1$, $\sigma_e = 0.1$	$R = 120$	0.016	0.040	0.002	0.023	0.000	0.021
	$R = 240$	0.022	0.040	0.004	0.033	0.001	0.023
	$R = 60$	0.011	0.033	0.002	0.025	0.000	0.031
$\phi = 0.1$, $\sigma_e = 1$	$R = 120$	0.013	0.033	0.003	0.029	0.000	0.028
	$R = 240$	0.026	0.042	0.009	0.030	0.001	0.031
	$R = 60$	0.010	0.039	0.001	0.029	0.000	0.027
$\phi = 0.5$, $\sigma_e = 0.1$	$R = 120$	0.016	0.041	0.003	0.032	0.001	0.025
	$R = 240$	0.022	0.037	0.006	0.035	0.002	0.030
	$R = 60$	0.011	0.037	0.001	0.024	0.000	0.021
$\phi = 0.5$, $\sigma_e = 1$	$R = 120$	0.019	0.043	0.004	0.027	0.001	0.021
	$R = 240$	0.019	0.036	0.007	0.027	0.002	0.023
	$R = 60$	0.012	0.053	0.001	0.036	0.000	0.025
$\phi = 0.9$, $\sigma_e = 0.1$	$R = 120$	0.019	0.044	0.002	0.029	0.001	0.027
	$R = 240$	0.020	0.038	0.006	0.034	0.001	0.027
	$R = 60$	0.007	0.037	0.001	0.026	0.000	0.020
$\phi = 0.9$, $\sigma_e = 1$	$R = 120$	0.017	0.040	0.003	0.027	0.000	0.025
	$R = 240$	0.025	0.039	0.007	0.026	0.002	0.022
	$R = 60$	0.010	0.047	0.000	0.038	0.000	0.032
$\phi = 0.95$, $\sigma_e = 0.1$	$R = 120$	0.015	0.043	0.003	0.032	0.001	0.024
	$R = 240$	0.021	0.041	0.008	0.031	0.001	0.024
	$R = 60$	0.010	0.033	0.001	0.024	0.000	0.024
$\phi = 0.95$, $\sigma_e = 1$	$R = 120$	0.012	0.028	0.003	0.028	0.000	0.025
	$R = 240$	0.017	0.035	0.007	0.031	0.003	0.028
	$R = 60$	0.008	0.083	0.001	0.062	0.000	0.039
$\phi = 0.99$, $\sigma_e = 0.1$	$R = 120$	0.017	0.058	0.002	0.043	0.000	0.028
	$R = 240$	0.019	0.044	0.004	0.030	0.001	0.027
	$R = 60$	0.005	0.029	0.000	0.024	0.000	0.030
$\phi = 0.99$, $\sigma_e = 1$	$R = 120$	0.015	0.042	0.002	0.030	0.001	0.027
	$R = 240$	0.022	0.039	0.005	0.028	0.001	0.024

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Recursive Scheme).

Table 11: Rejection frequency at 5% level, $b = 0.1$ (With intercept on Model 1)

Repeat = 2000		$P = 48$ DM_P	$P = 48$ ENC_P	$P = 240$ DM_P	$P = 240$ ENC_P	$P = 1200$ DM_P	$P = 1200$ ENC_P
	$R = 60$	0.943	1.000	1.000	1.000	1.000	1.000
$\phi = 0$, $\sigma_e = 0.1$	$R = 120$	0.938	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.947	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.034	0.102	0.039	0.293	0.245	0.883
$\phi = 0$, $\sigma_e = 1$	$R = 120$	0.050	0.117	0.066	0.324	0.321	0.901
	$R = 240$	0.070	0.147	0.096	0.367	0.365	0.907
	$R = 60$	0.934	1.000	1.000	1.000	1.000	1.000
$\phi = 0.1$, $\sigma_e = 0.1$	$R = 120$	0.950	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.944	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.033	0.102	0.046	0.282	0.278	0.878
$\phi = 0.1$, $\sigma_e = 1$	$R = 120$	0.053	0.130	0.058	0.315	0.340	0.902
	$R = 240$	0.074	0.140	0.107	0.365	0.388	0.926
	$R = 60$	0.967	1.000	1.000	1.000	1.000	1.000
$\phi = 0.5$, $\sigma_e = 0.1$	$R = 120$	0.968	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.983	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.047	0.135	0.058	0.378	0.410	0.941
$\phi = 0.5$, $\sigma_e = 1$	$R = 120$	0.055	0.144	0.089	0.415	0.485	0.966
	$R = 240$	0.078	0.166	0.140	0.466	0.525	0.971
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.9$, $\sigma_e = 0.1$	$R = 120$	0.999	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.116	0.359	0.394	0.901	0.995	1.000
$\phi = 0.9$, $\sigma_e = 1$	$R = 120$	0.159	0.388	0.446	0.920	0.995	1.000
	$R = 240$	0.175	0.427	0.508	0.940	0.997	1.000
	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.95$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.170	0.490	0.675	0.979	1.000	1.000
$\phi = 0.95$, $\sigma_e = 1$	$R = 120$	0.237	0.569	0.724	0.989	1.000	1.000
	$R = 240$	0.279	0.603	0.758	0.992	1.000	1.000
	$R = 60$	0.999	1.000	1.000	1.000	1.000	1.000
$\phi = 0.99$, $\sigma_e = 0.1$	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 60$	0.331	0.687	0.951	0.999	1.000	1.000
$\phi = 0.99$, $\sigma_e = 1$	$R = 120$	0.459	0.775	0.971	1.000	1.000	1.000
	$R = 240$	0.534	0.835	0.962	1.000	1.000	1.000

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Recursive Scheme).

Table 12: Rejection frequency at 5% level, $b = 1$ (With intercept on Model 1)

$Repeat = 2000$		$P = 48$ DM_P	ENC_P	$P = 240$ DM_P	ENC_P	$P = 1200$ DM_P	ENC_P
$\phi = 0, \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0, \sigma_e = 1$	$R = 60$	0.950	1.000	1.000	1.000	1.000	1.000
	$R = 120$	0.955	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.948	1.000	1.000	1.000	1.000	1.000
$\phi = 0.1, \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.1, \sigma_e = 1$	$R = 60$	0.947	1.000	1.000	1.000	1.000	1.000
	$R = 120$	0.936	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.956	1.000	1.000	1.000	1.000	1.000
$\phi = 0.5, \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.5, \sigma_e = 1$	$R = 60$	0.980	1.000	1.000	1.000	1.000	1.000
	$R = 120$	0.975	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.974	1.000	1.000	1.000	1.000	1.000
$\phi = 0.9, \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.9, \sigma_e = 1$	$R = 60$	0.999	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	0.999	1.000	1.000	1.000	1.000	1.000
$\phi = 0.95, \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.95, \sigma_e = 1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.99, \sigma_e = 0.1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000
$\phi = 0.99, \sigma_e = 1$	$R = 60$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 120$	1.000	1.000	1.000	1.000	1.000	1.000
	$R = 240$	1.000	1.000	1.000	1.000	1.000	1.000

Notes: The table shows the size of DM_P and ENC_P test under 5% nominal size from Monte-Carlo Simulation of 2000 times (Recursive Scheme).

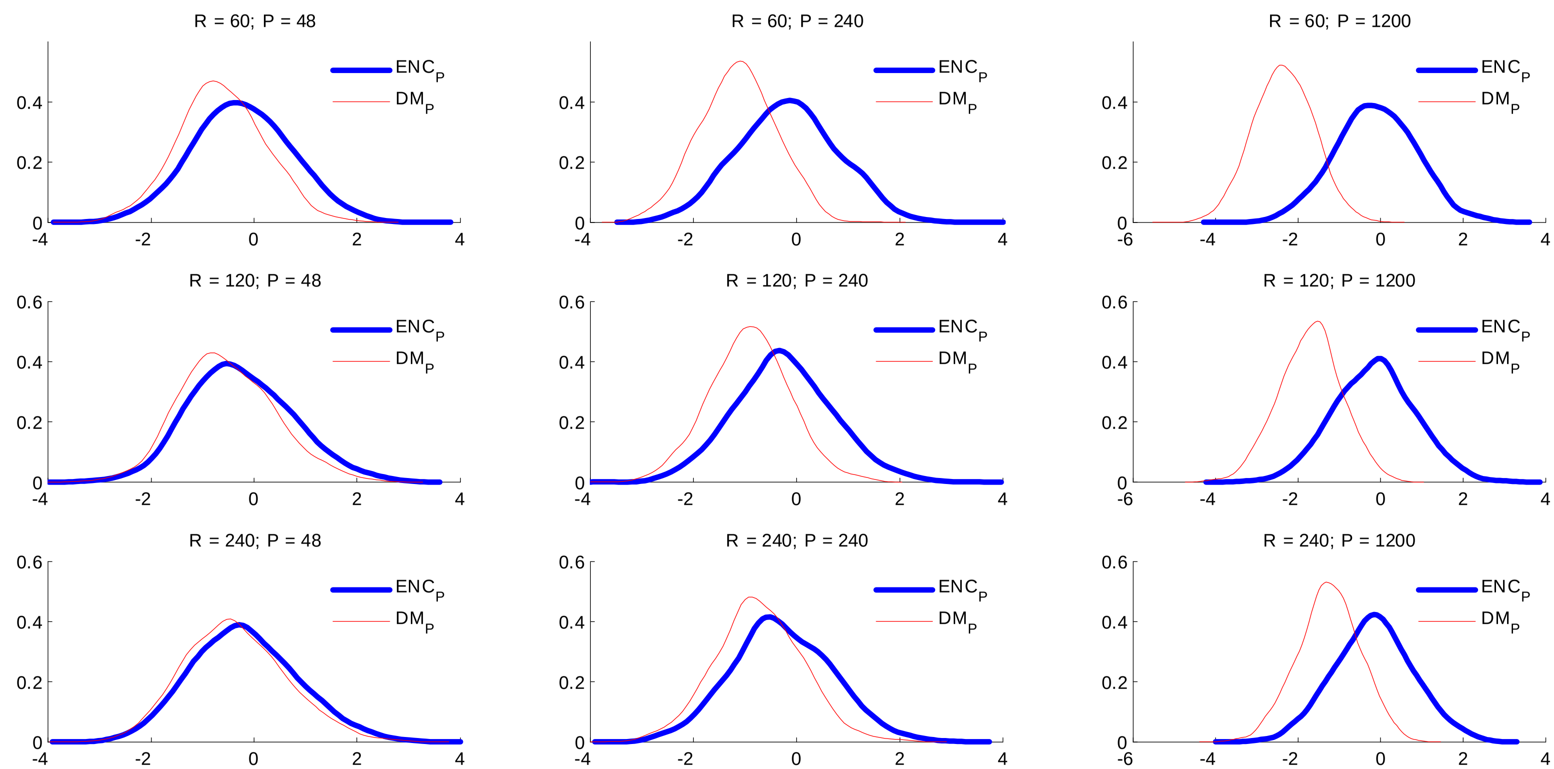

Figure 1: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0$, $b = 0$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats.

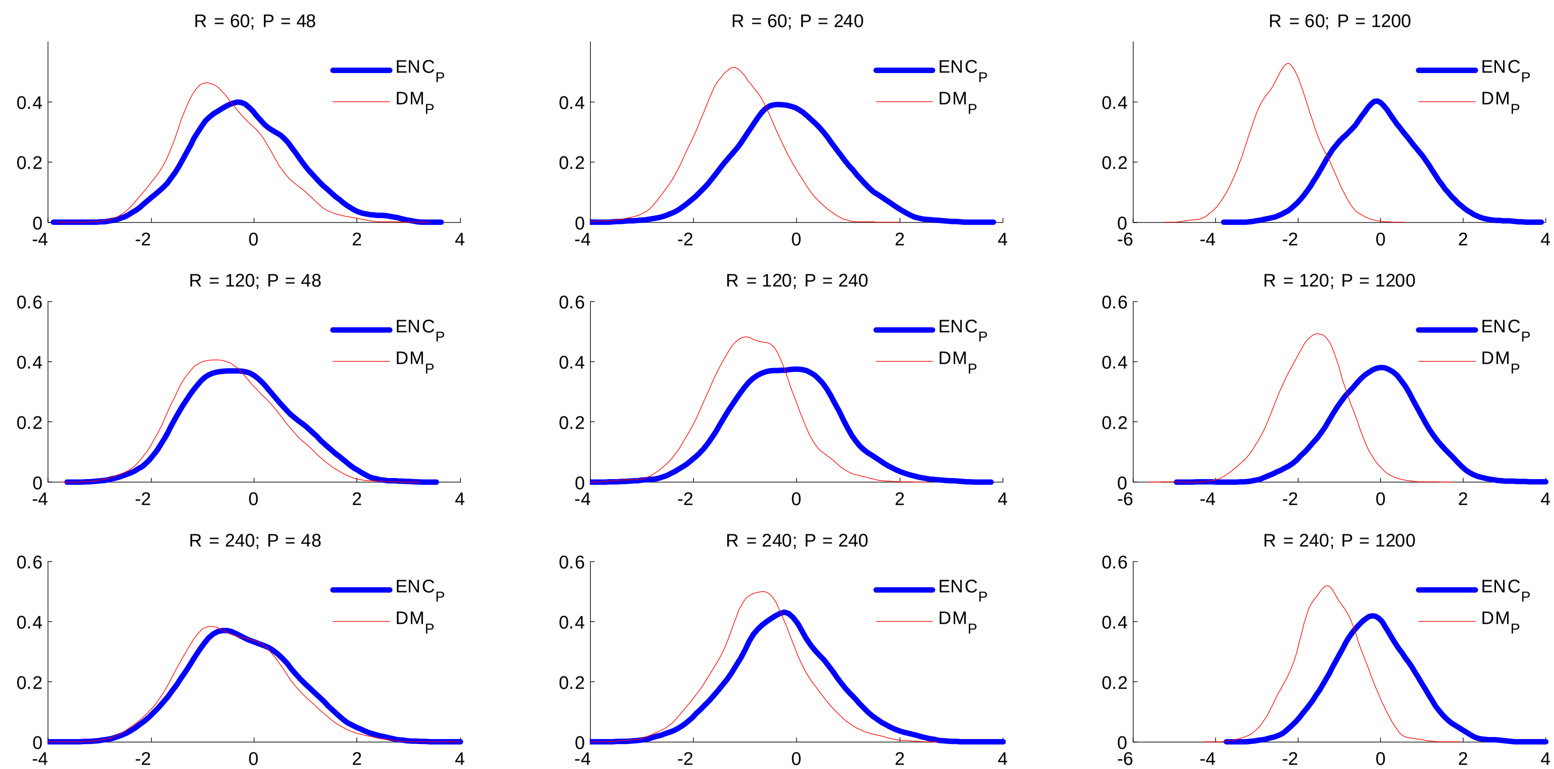

Figure 2: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0$, $b = 0$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats.

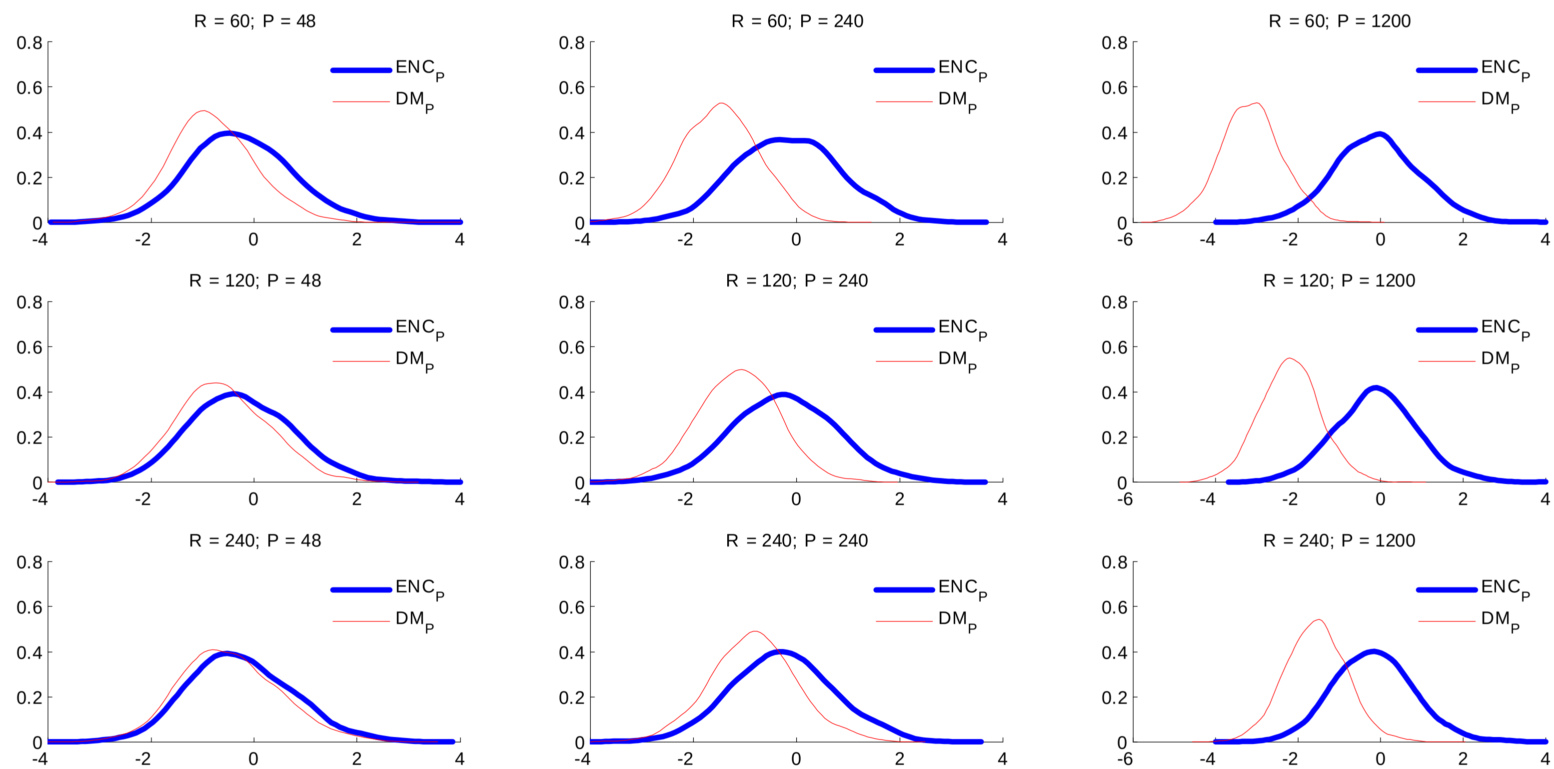

Figure 3: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0.99$, $b = 0$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats.

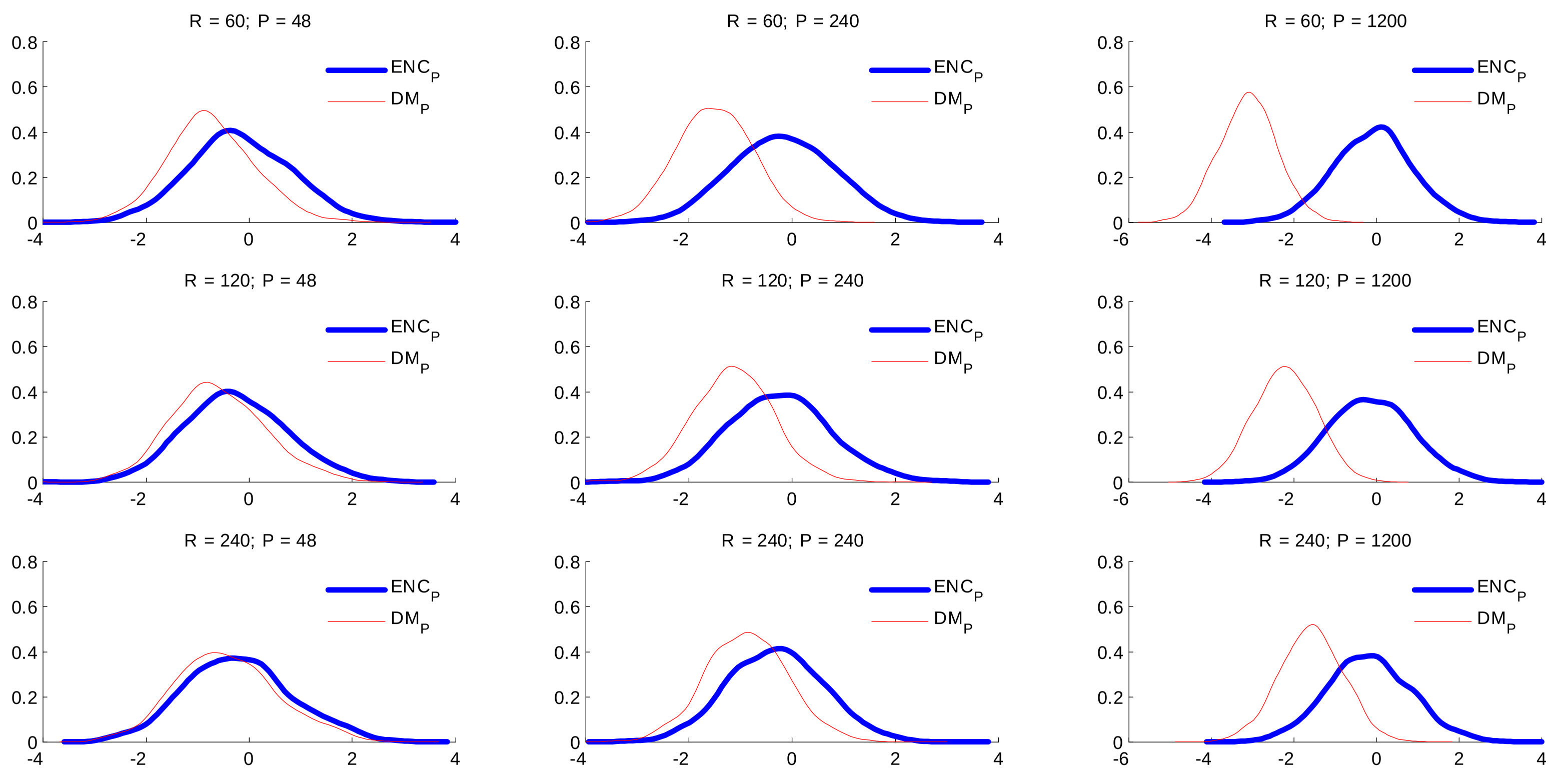

Figure 4: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0.99$, $b = 0$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats.

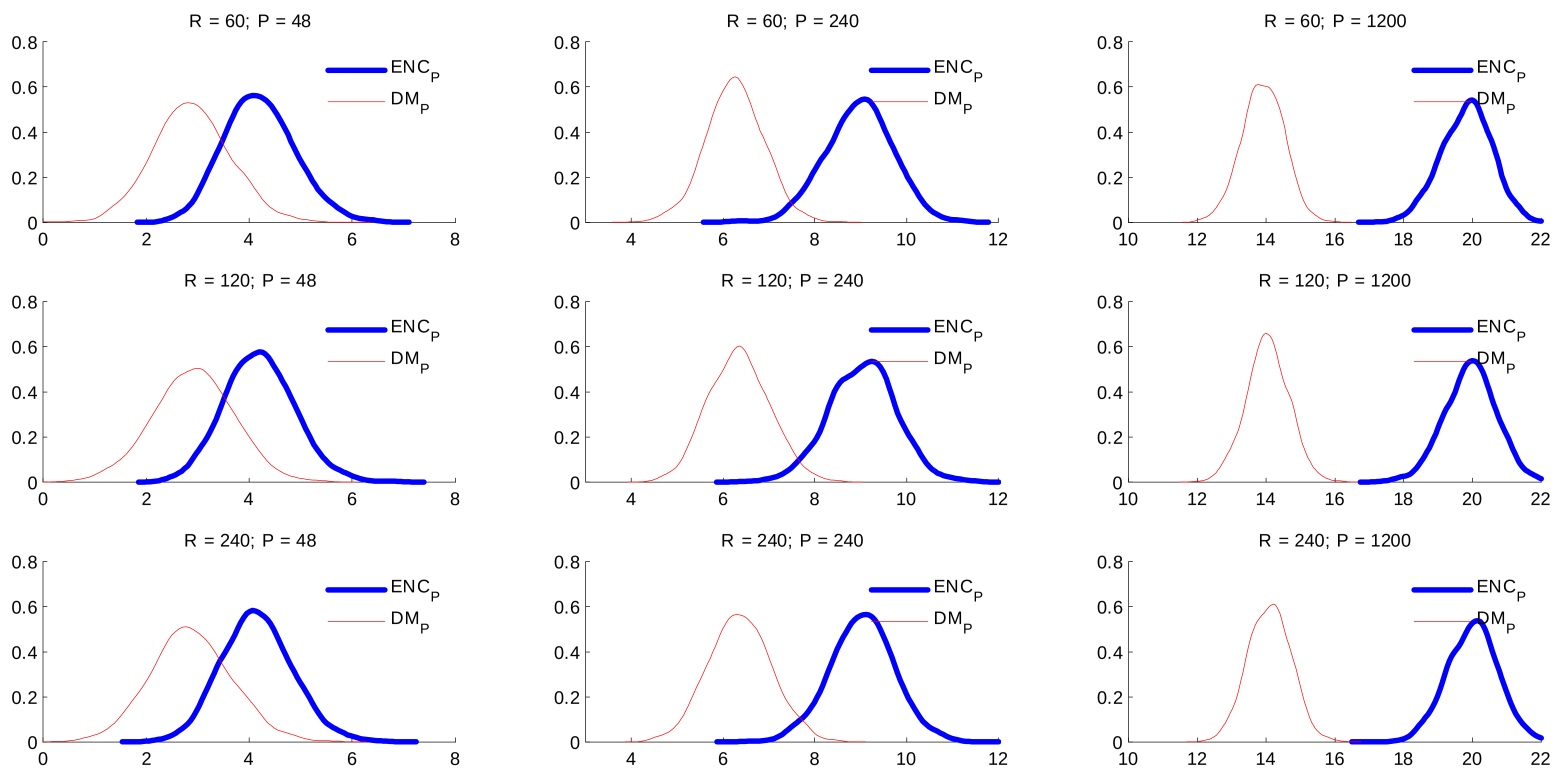

Figure 5: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 0.1$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats.

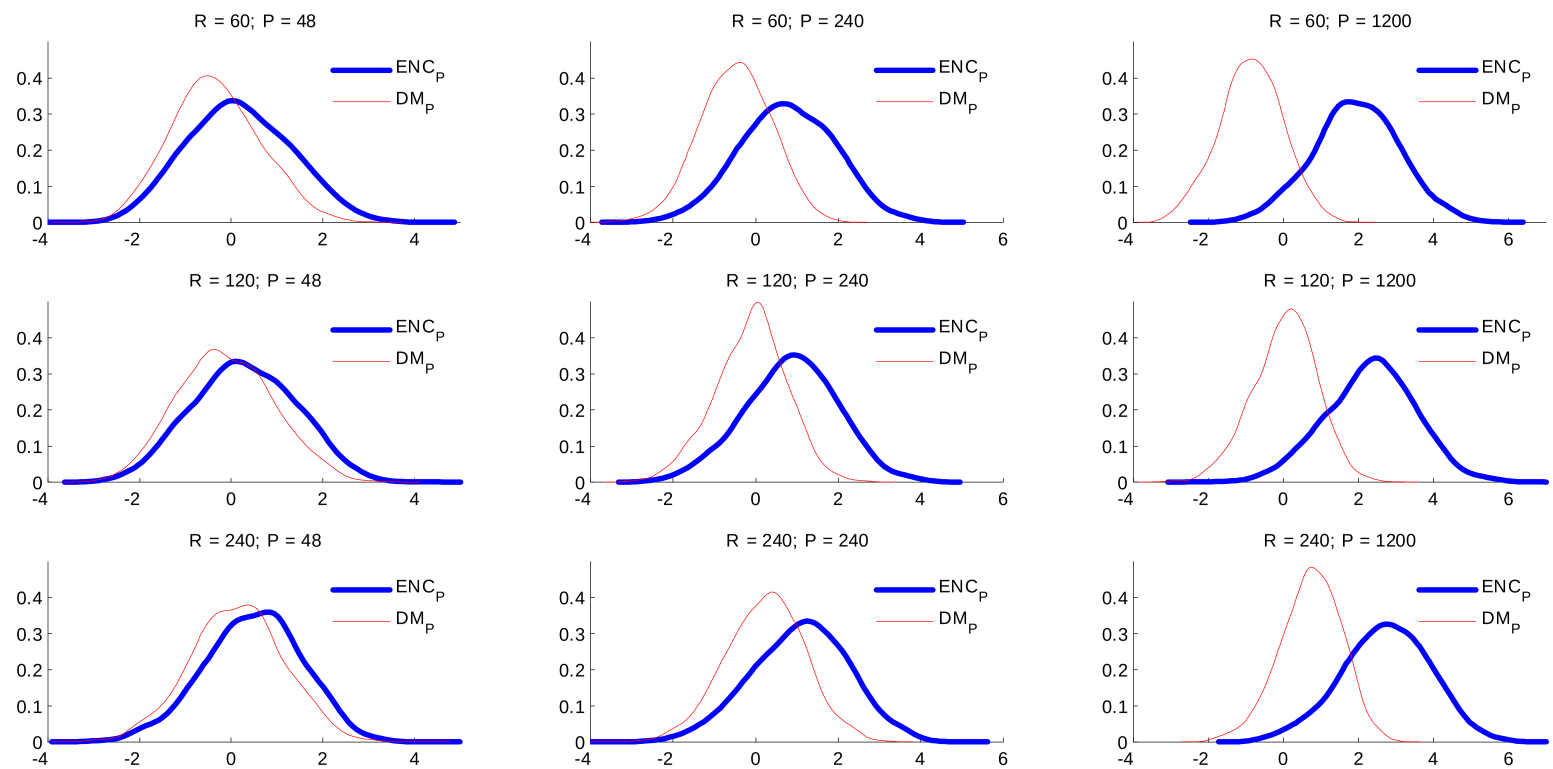

Figure 6: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 0.1$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats.

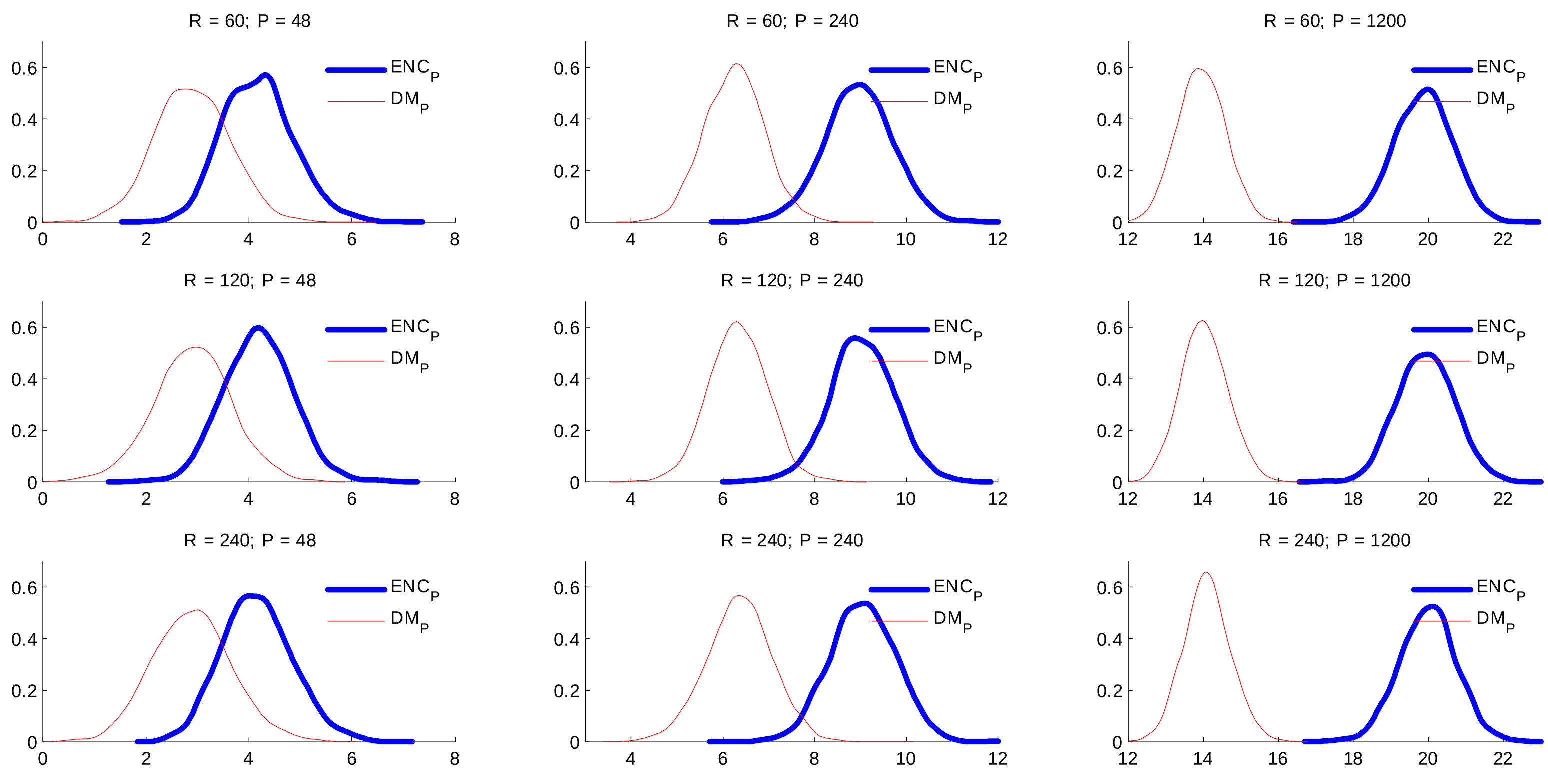

Figure 7: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 1$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats.

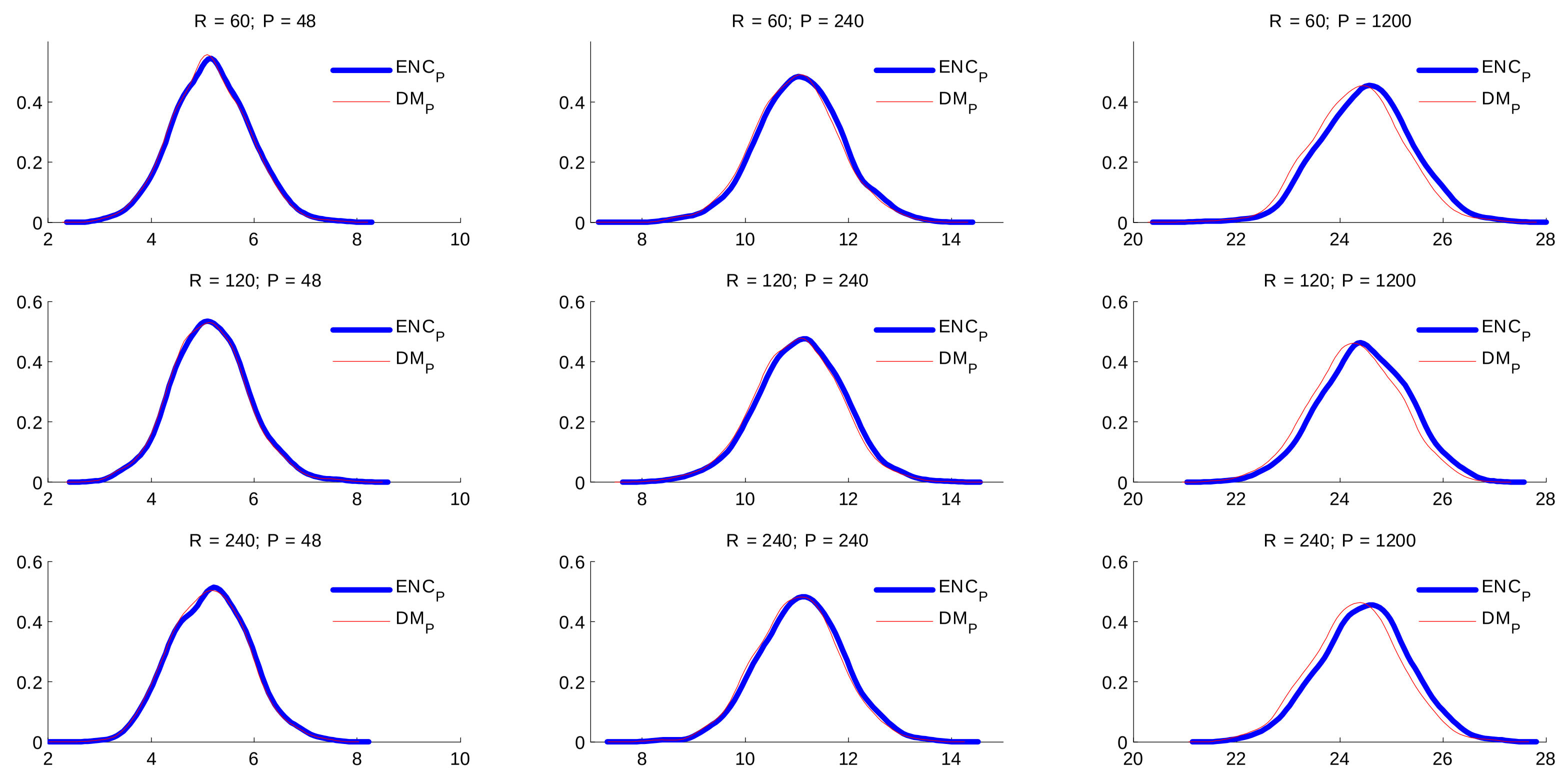

Figure 8: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 1$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats.

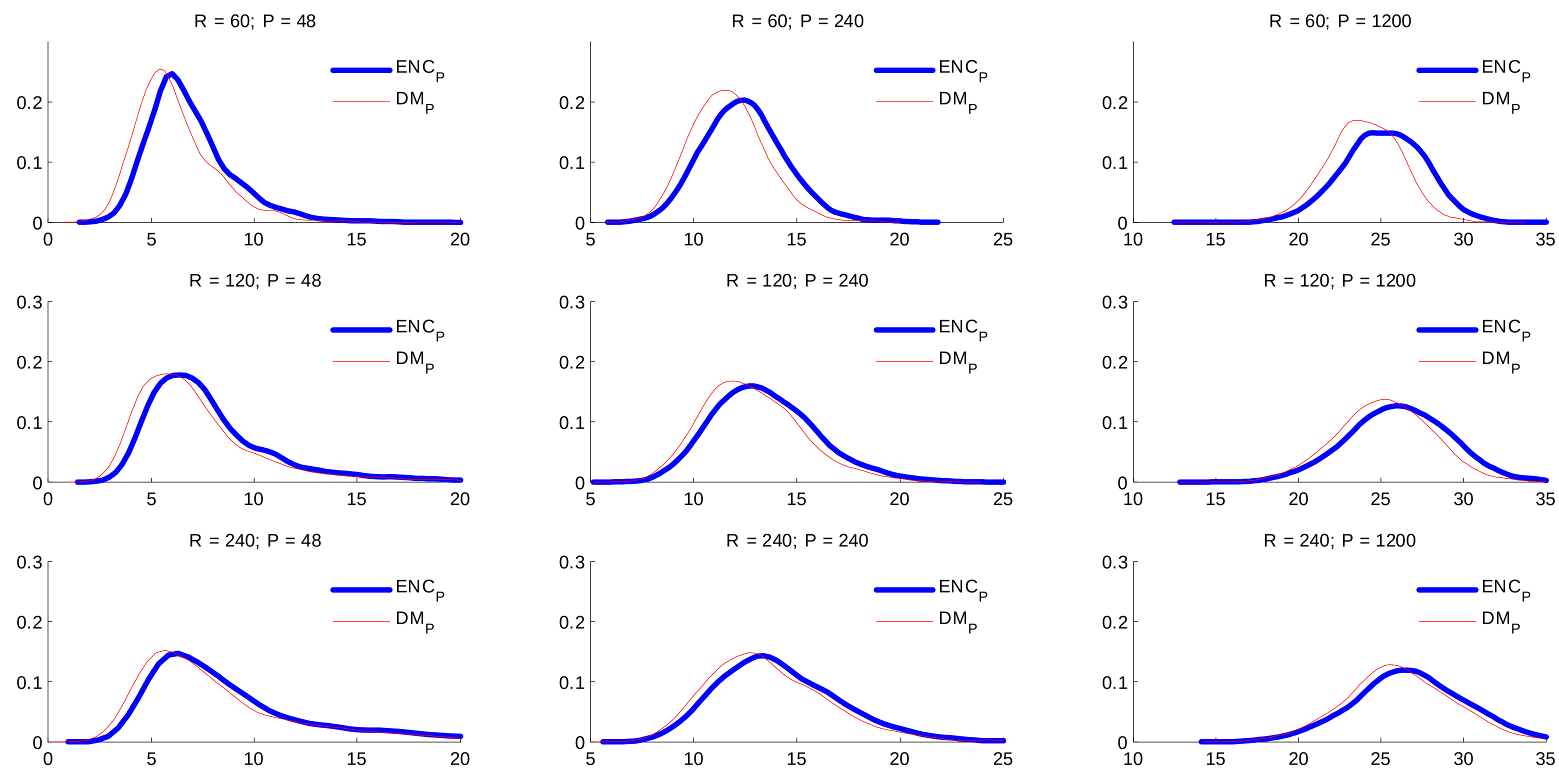

Figure 9: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 0.1$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats.

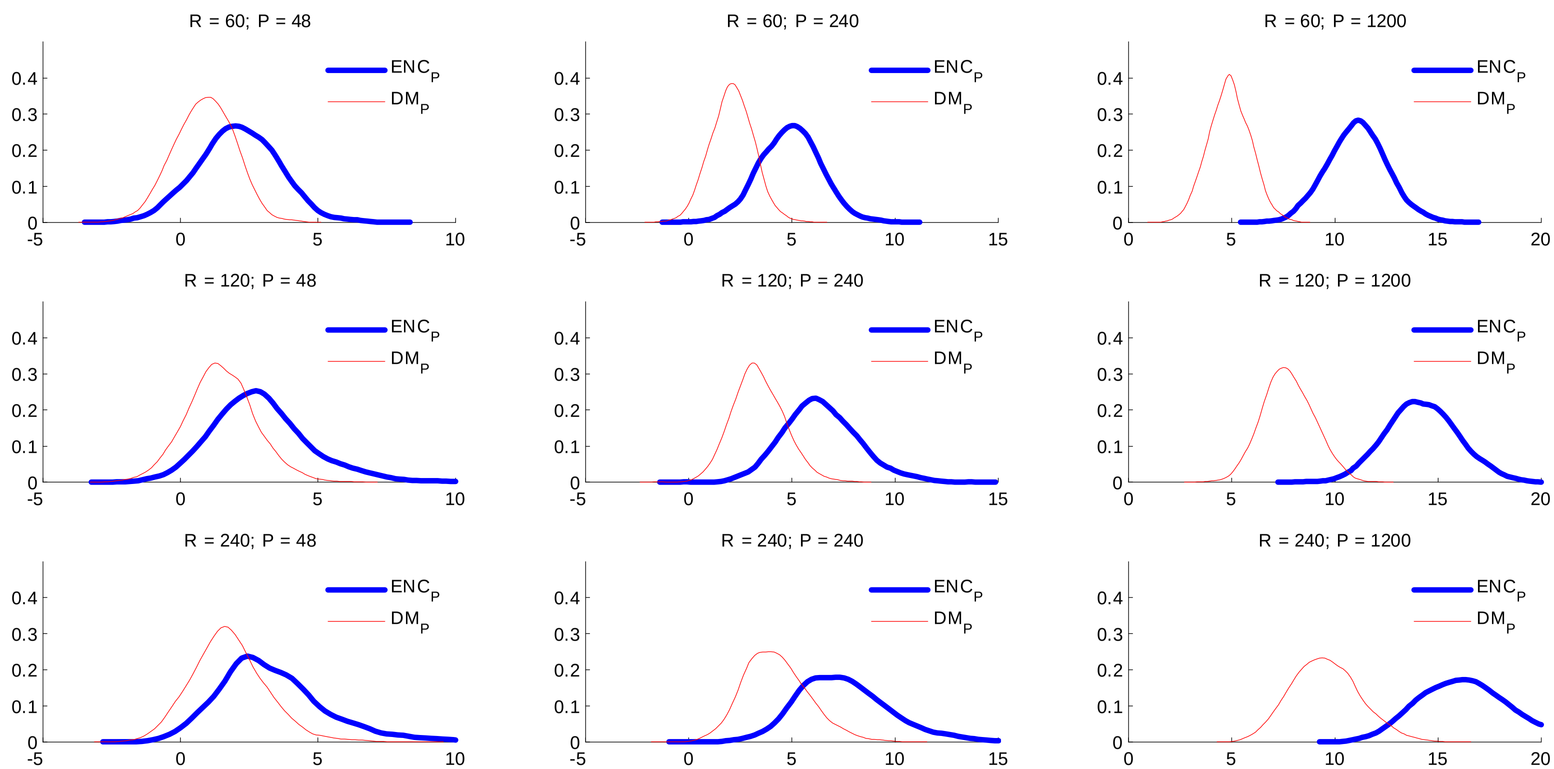

Figure 10: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 0.1$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats.

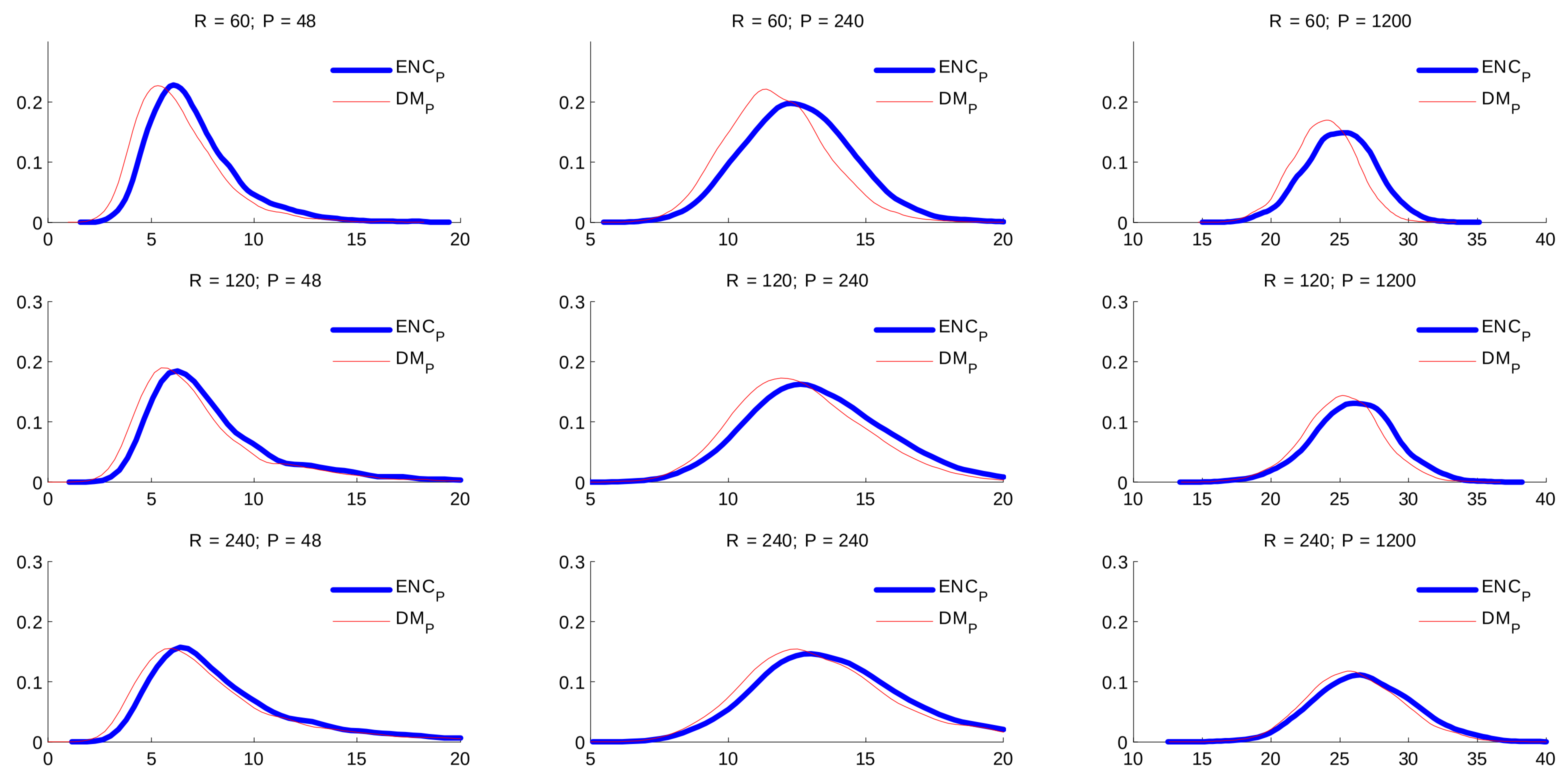

Figure 11: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 1$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats.

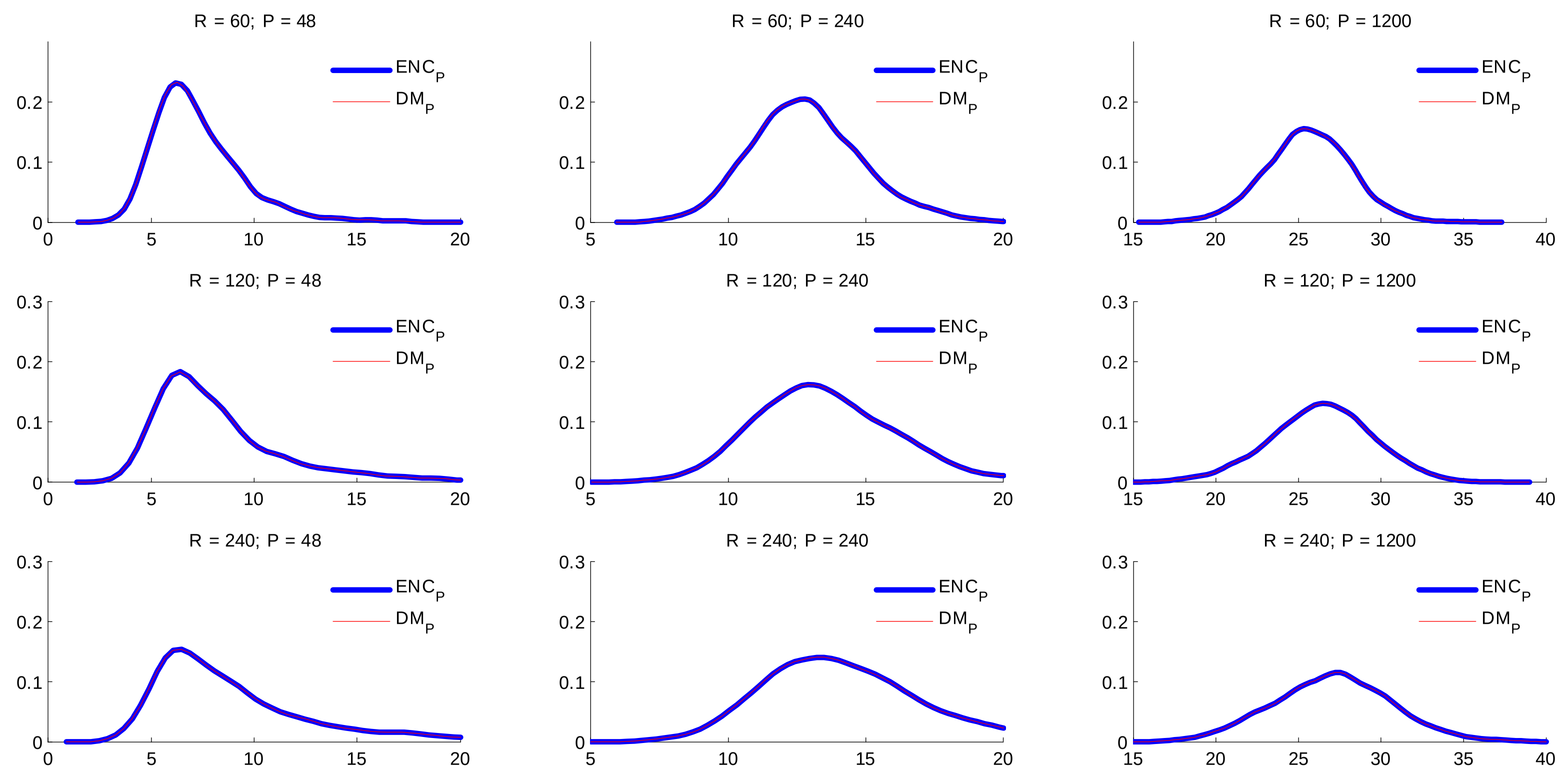

Figure 12: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 1$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats.

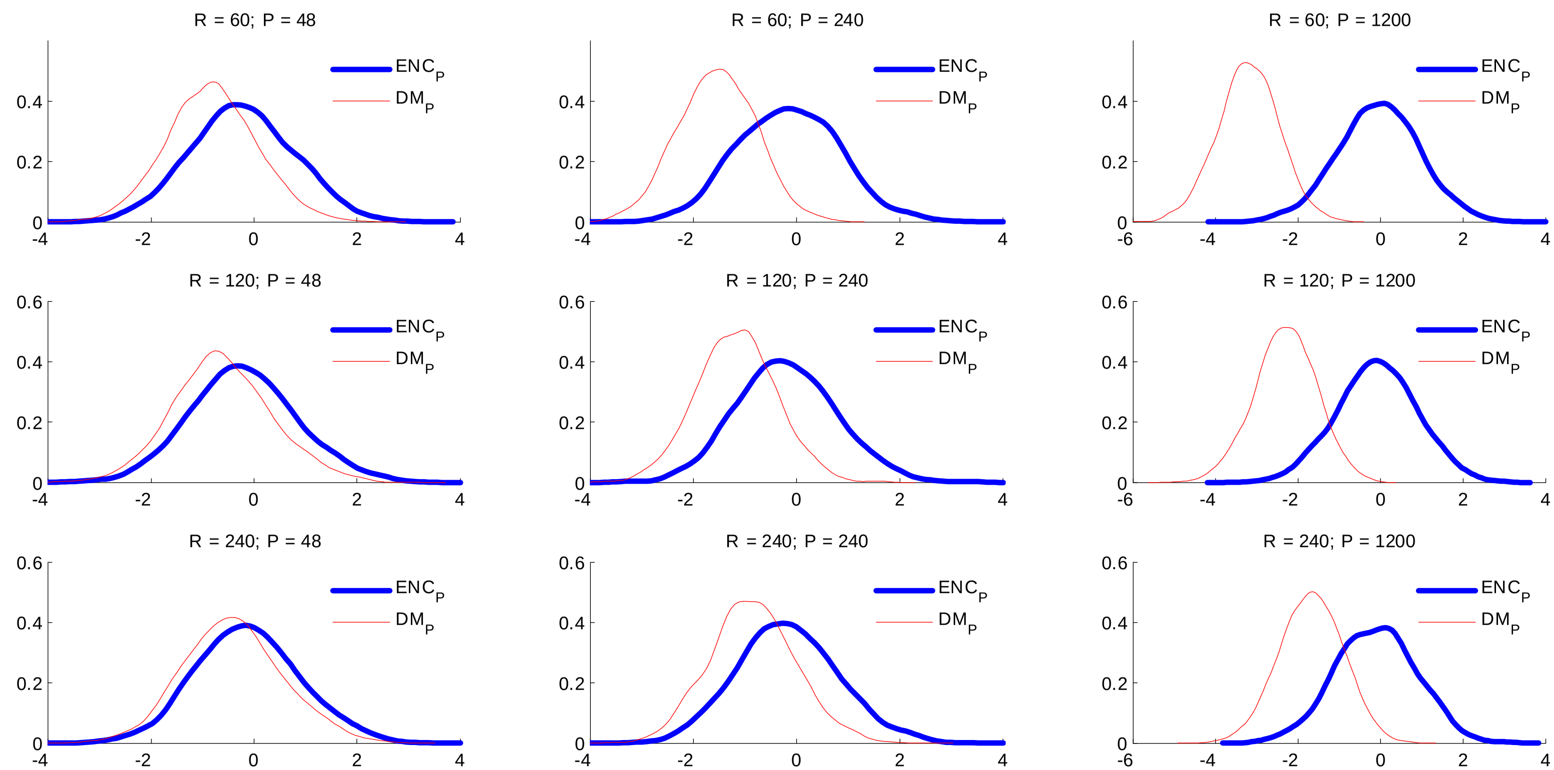

Figure 13: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0$, $b = 0$, $\sigma_e = 1$, without intercept on Model 1, 2000 Repeats.

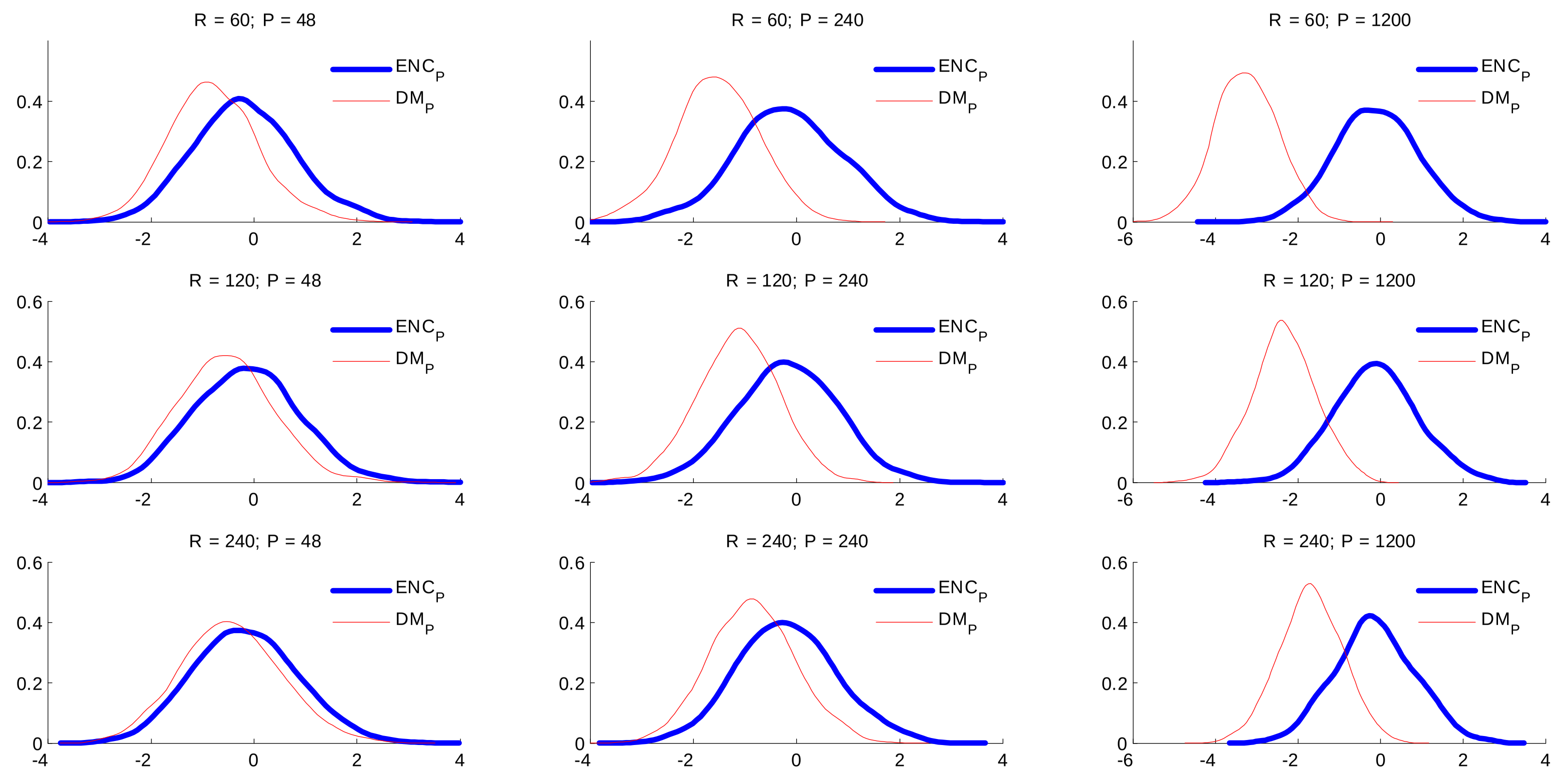

Figure 14: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0$, $b = 0$, $\sigma_e = 0.1$, without intercept on Model 1, 2000 Repeats.

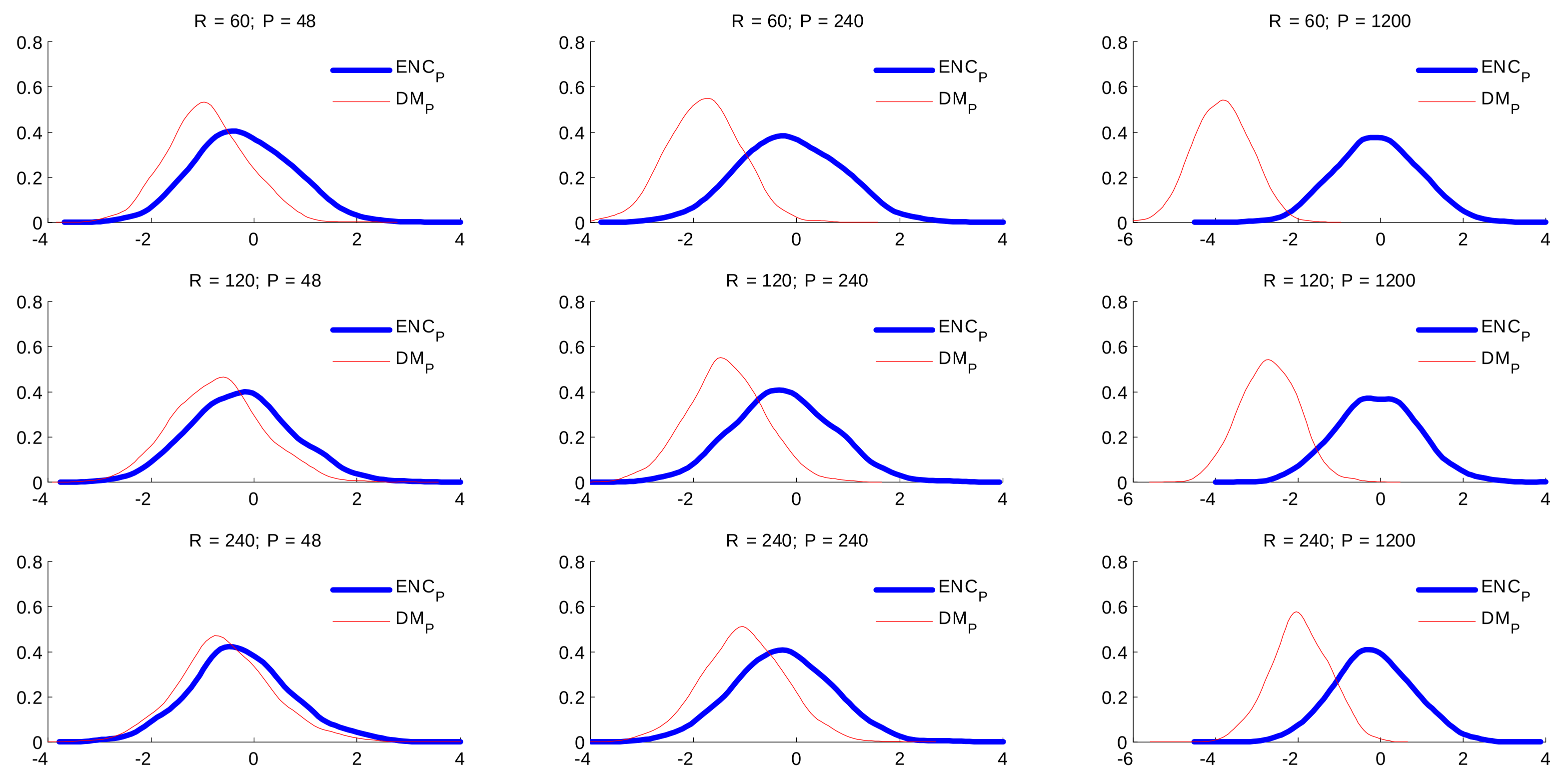

Figure 15: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0.99$, $b = 0$, $\sigma_e = 1$, without intercept on Model 1, 2000 Repeats.

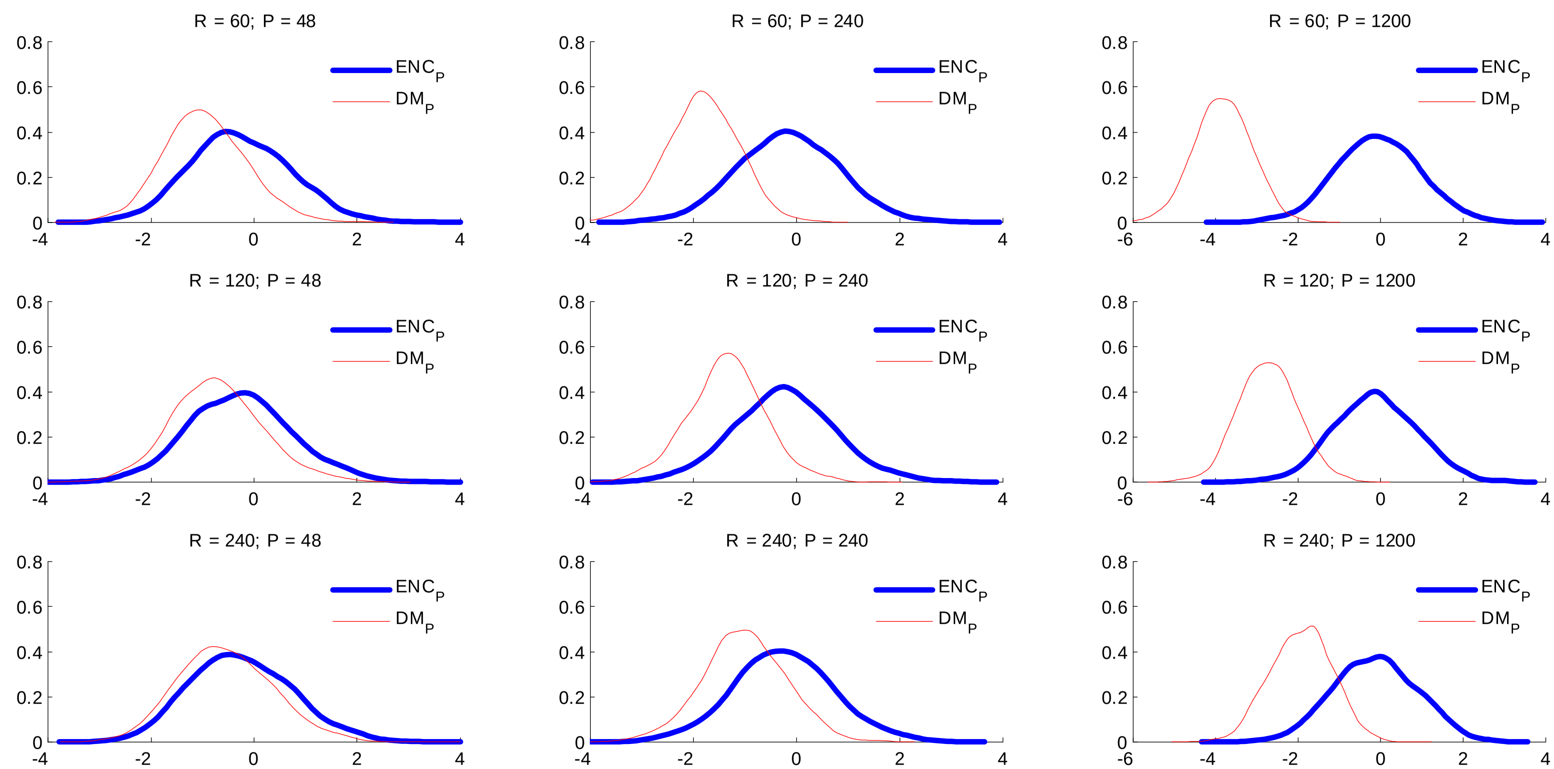

Figure 16: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0.99$, $b = 0$, $\sigma_e = 0.1$, without intercept on Model 1, 2000 Repeats.

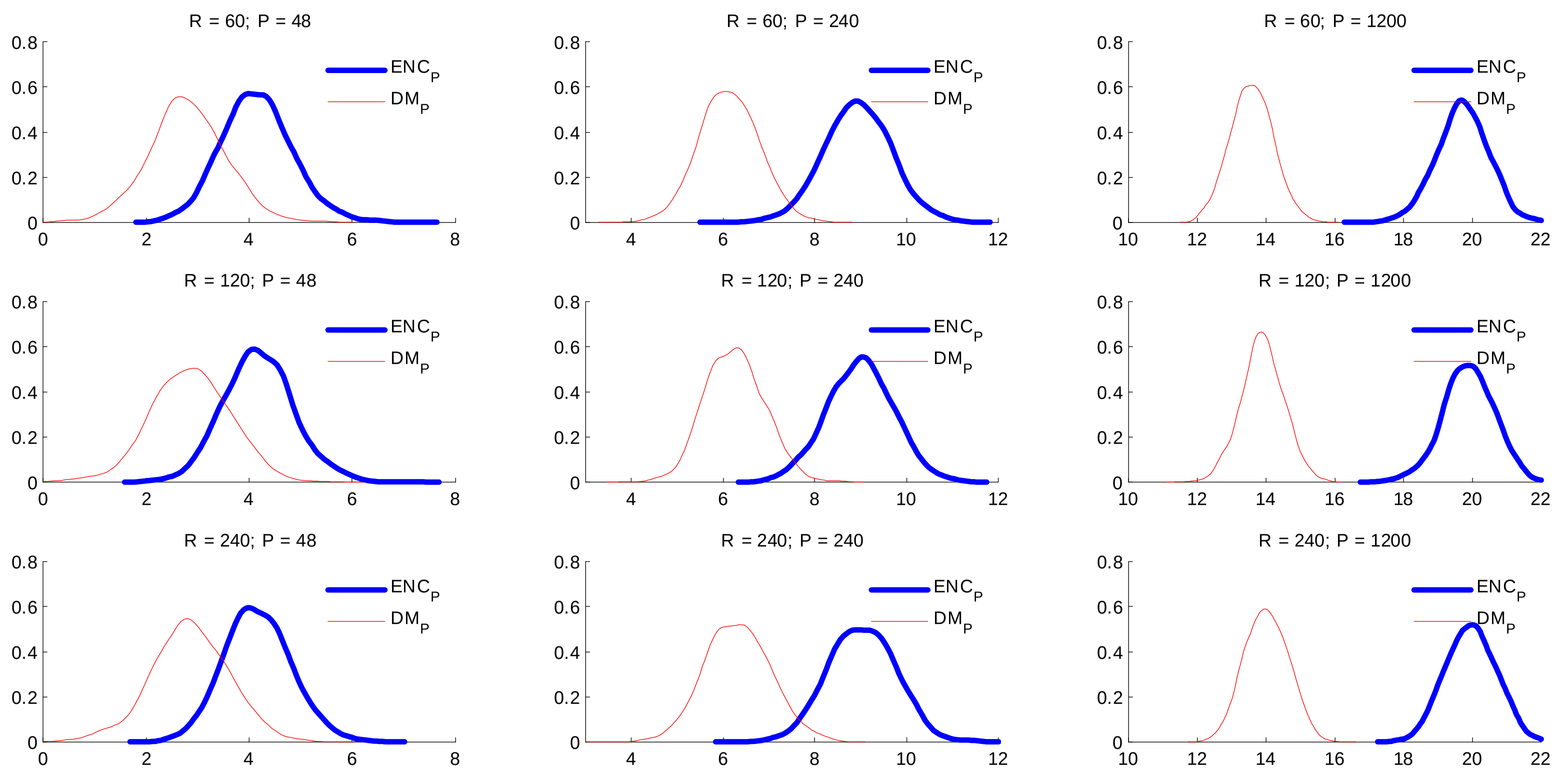

Figure 17: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 0.1$, $\sigma_e = 0.1$, without intercept on Model 1, 2000 Repeats.

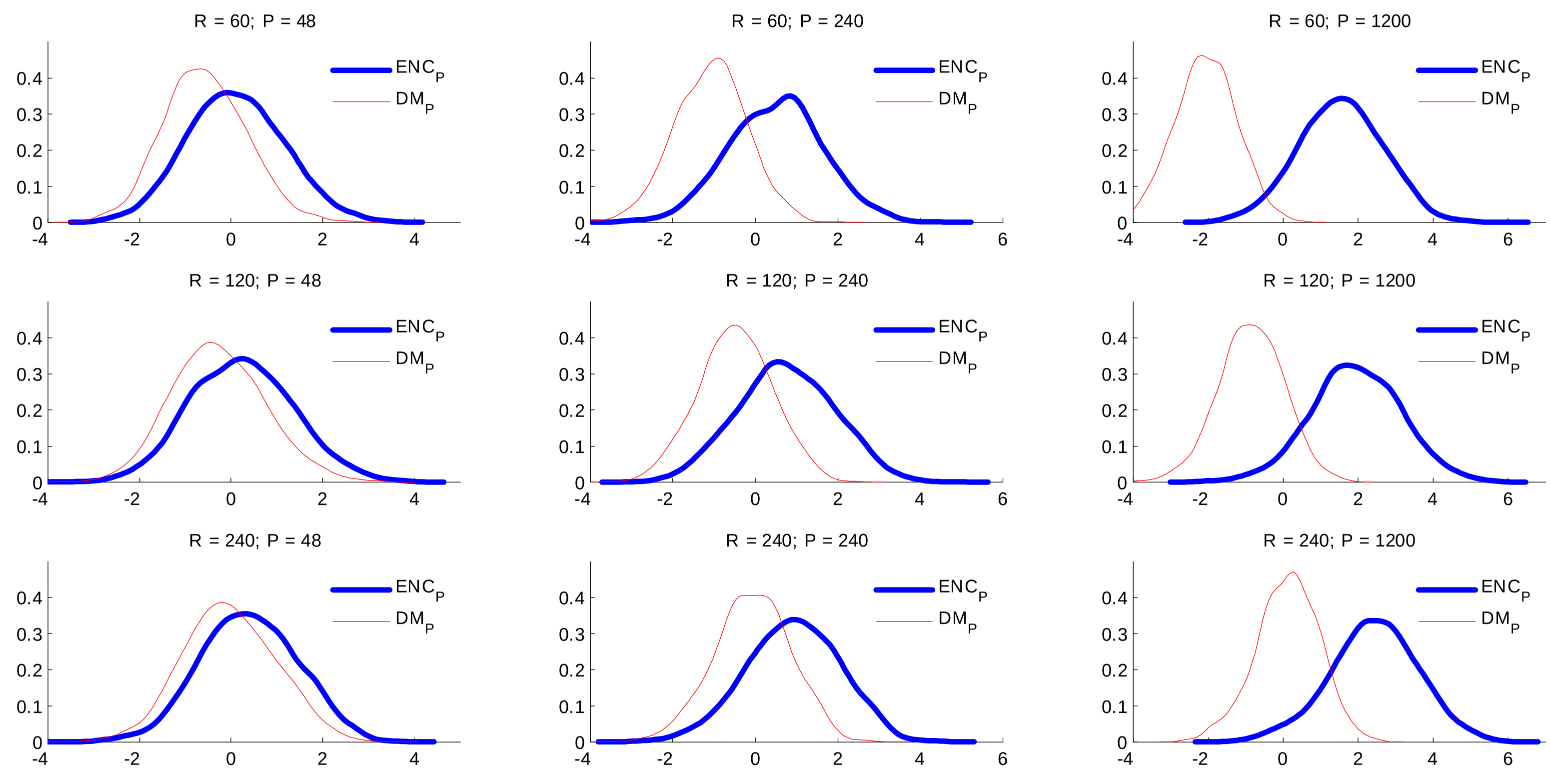

Figure 18: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 0.1$, $\sigma_e = 1$, without intercept on Model 1, 2000 Repeats.

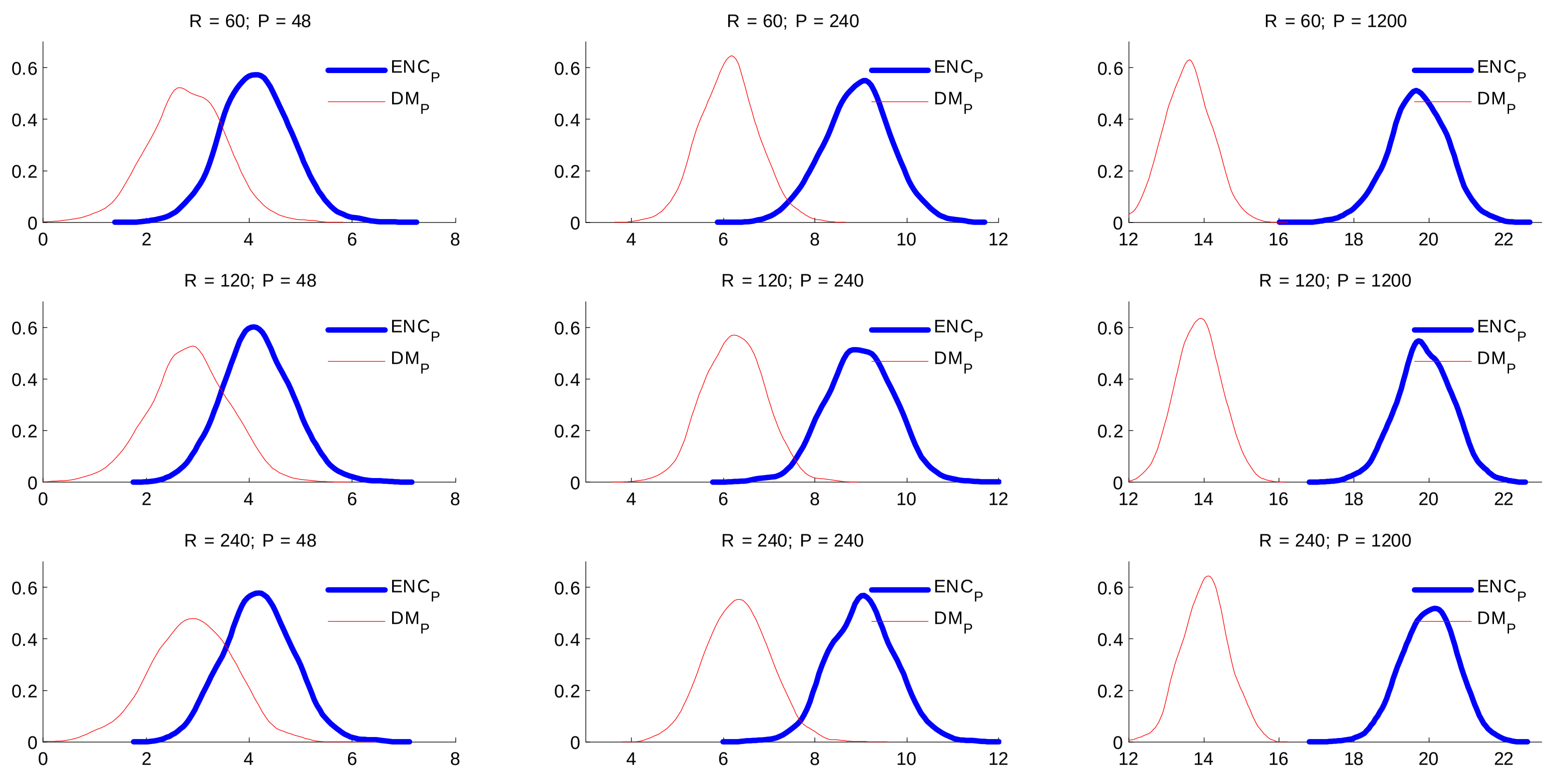

Figure 19: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 1$, $\sigma_e = 1$, without intercept on Model 1, 2000 Repeats.

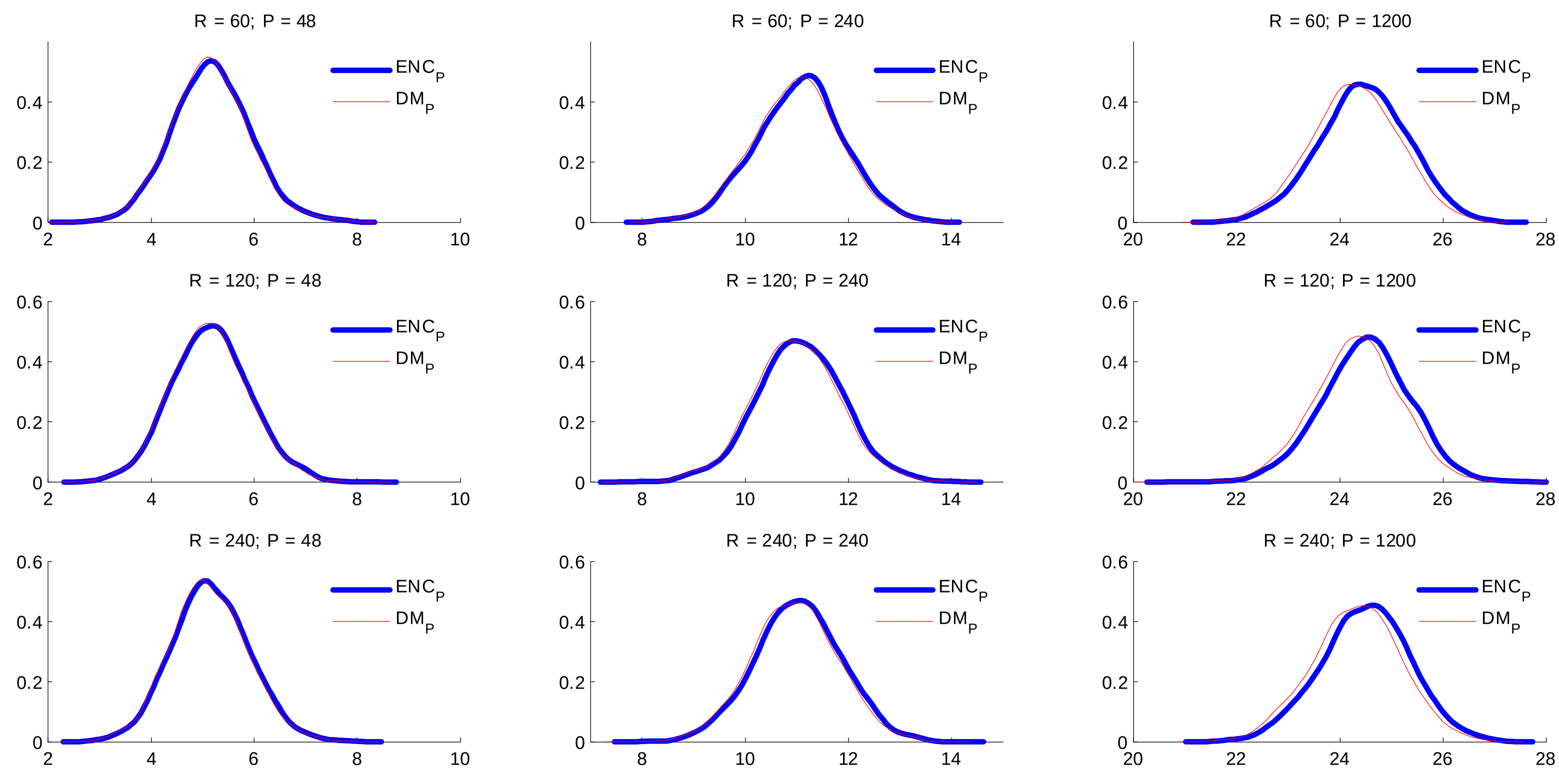

Figure 20: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 1$, $\sigma_e = 0.1$, without intercept on Model 1, 2000 Repeats.

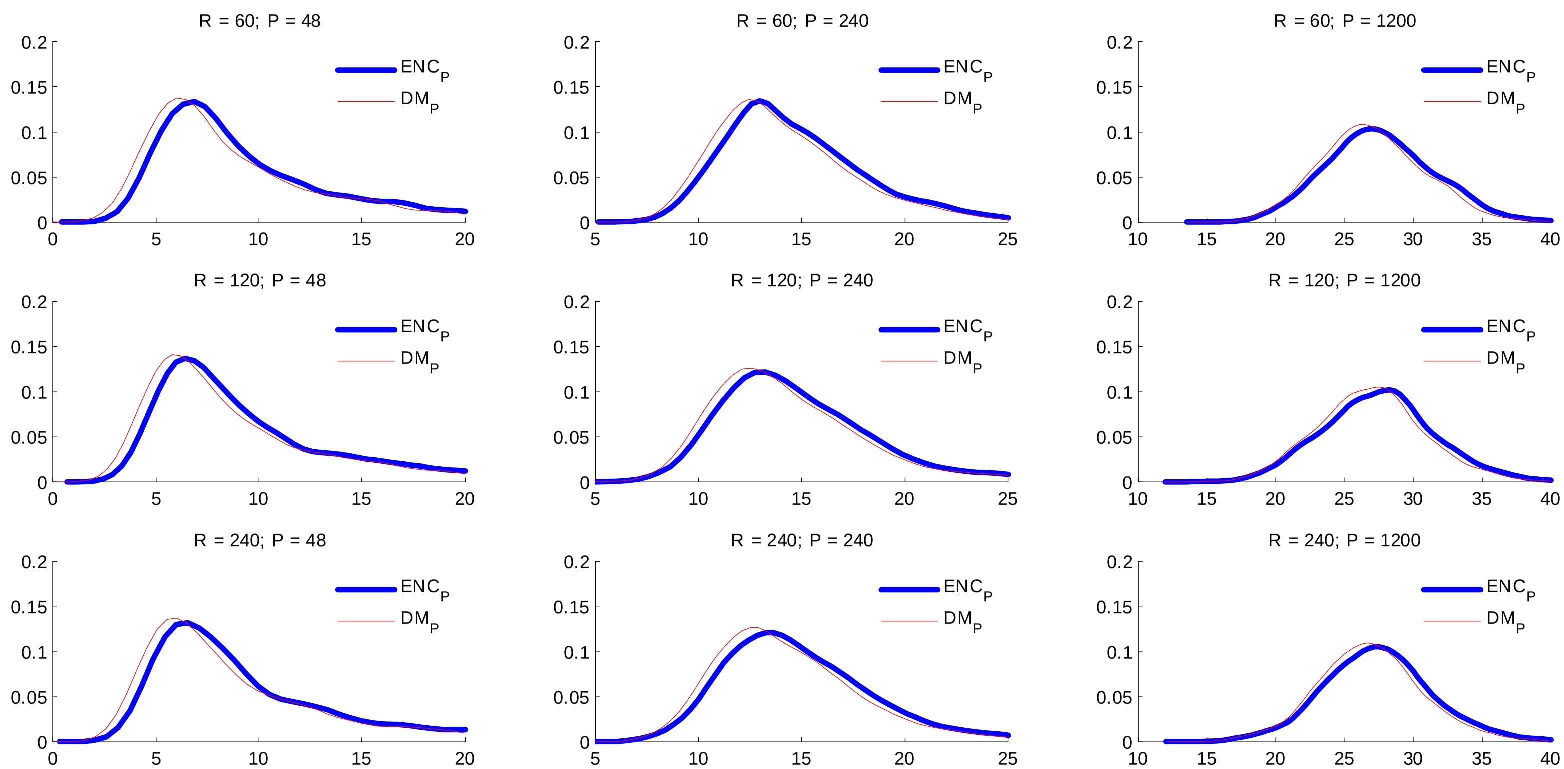

Figure 21: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 0.1$, $\sigma_e = 0.1$, without intercept on small model, 2000 Repeats.

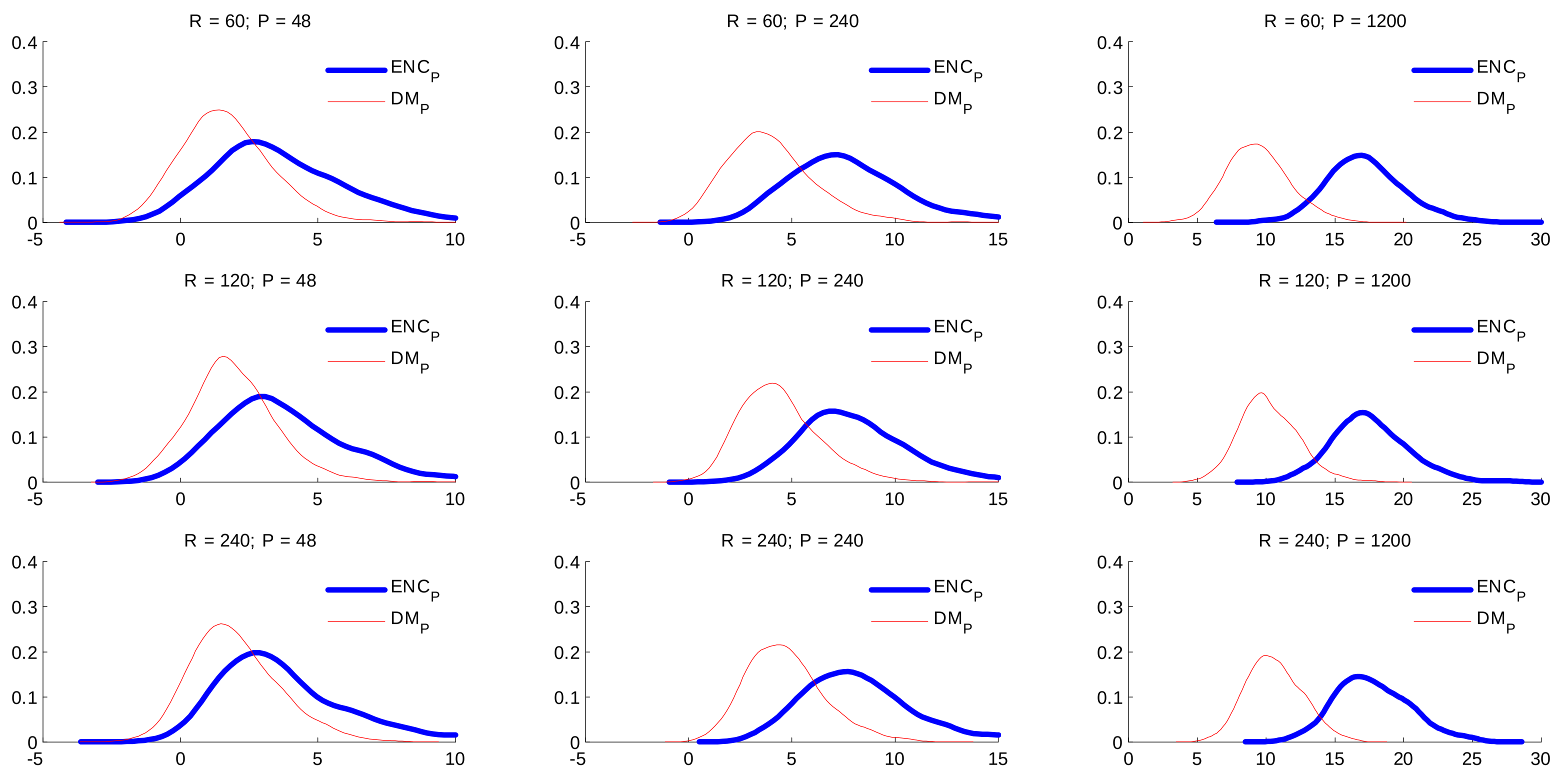

Figure 22: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 0.1$, $\sigma_e = 1$, without intercept on Model 1, 2000 Repeats.

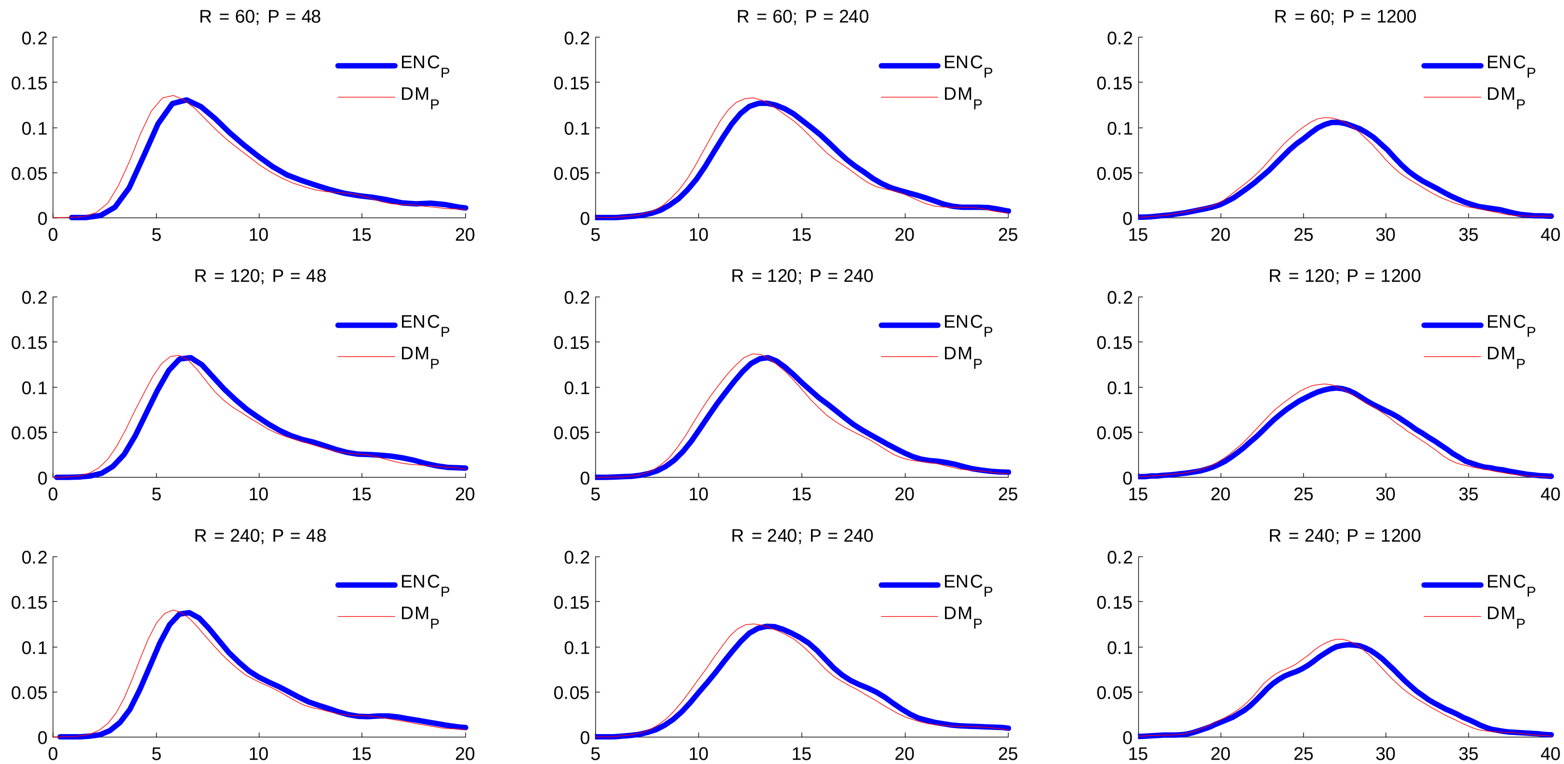

Figure 23: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 1$, $\sigma_e = 1$, without intercept on Model 1, 2000 Repeats.

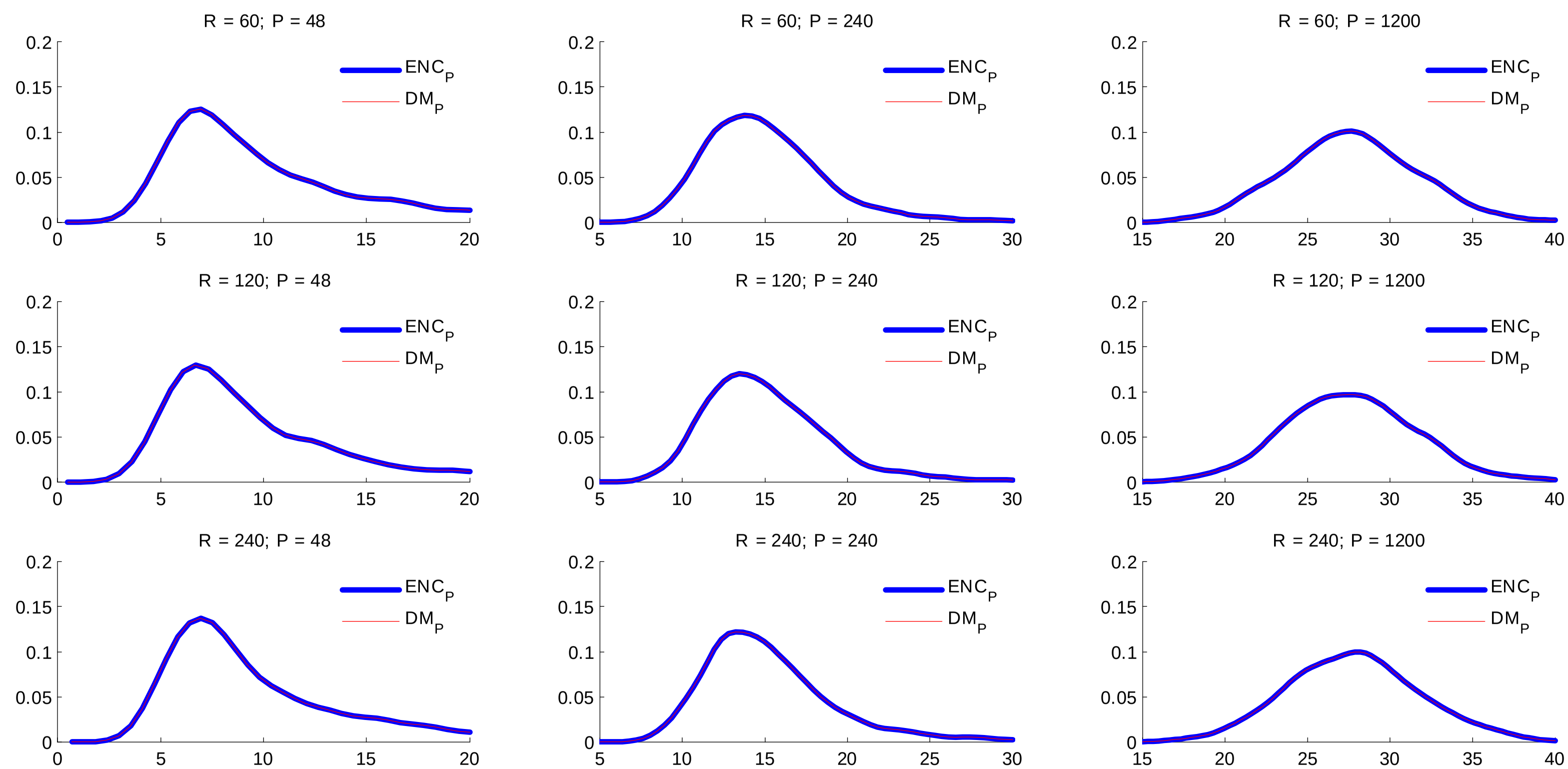

Figure 24: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 1$, $\sigma_e = 0.1$, without intercept on Model 1, 2000 Repeats.

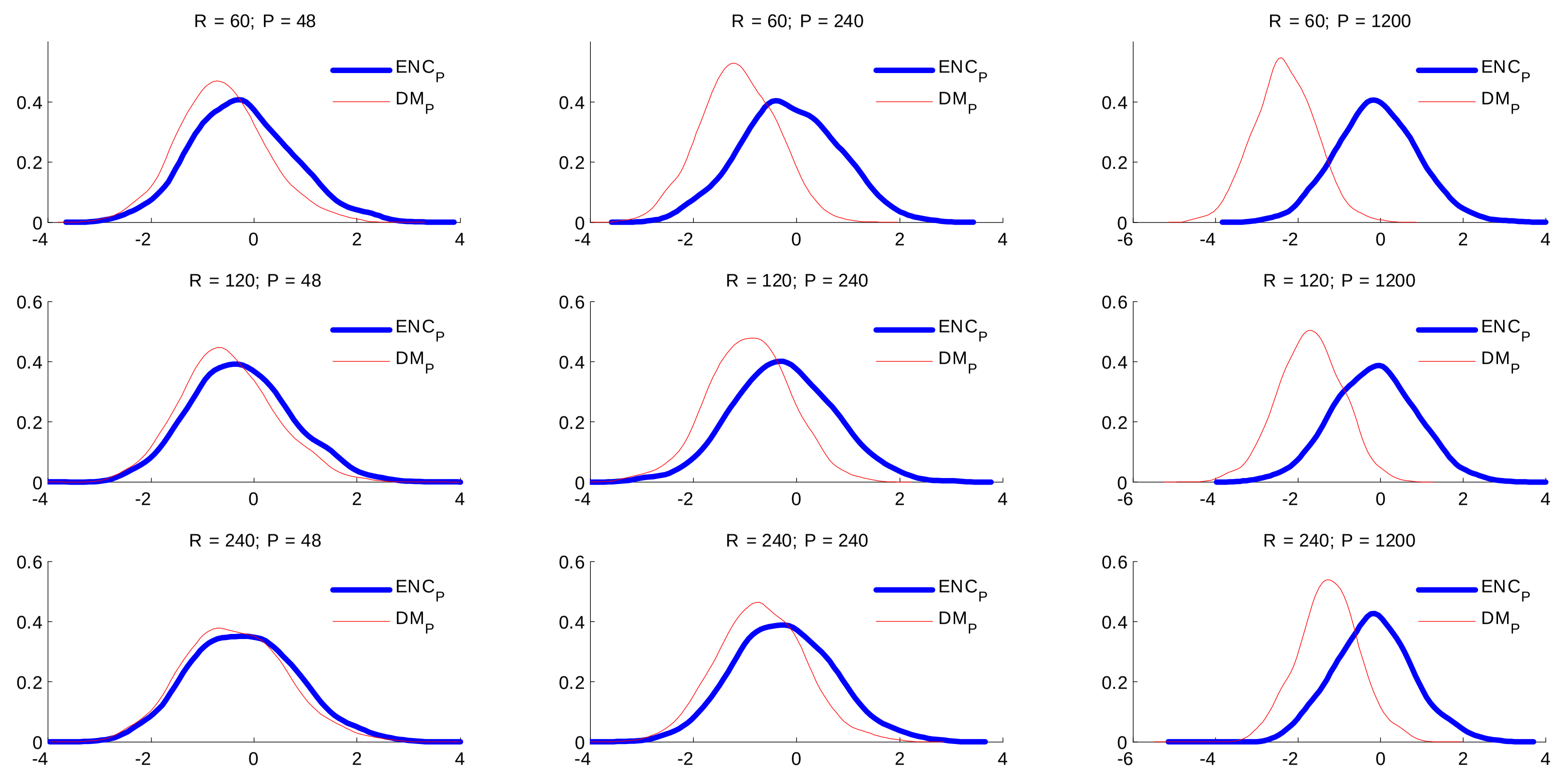

Figure 25: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0$, $b = 0$, $\sigma_e = 1$, $\rho(e_t, v_t) = 0.5$, with intercept on Model 1, 2000 Repeats.

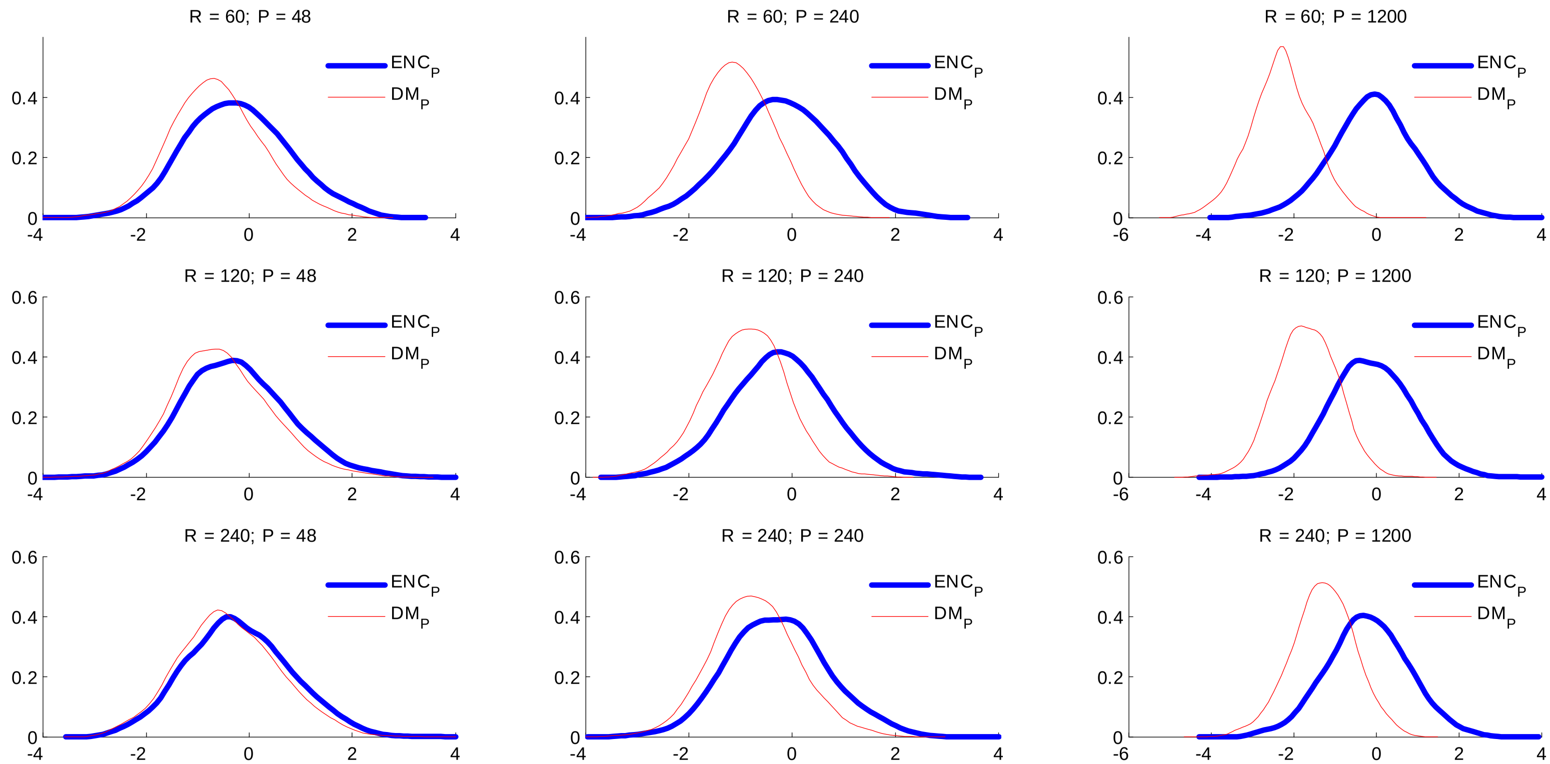

Figure 26: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0$, $b = 0$, $\sigma_e = 0.1$, $\rho(e_t, v_t) = 0.5$, with intercept on Model 1, 2000 Repeats.

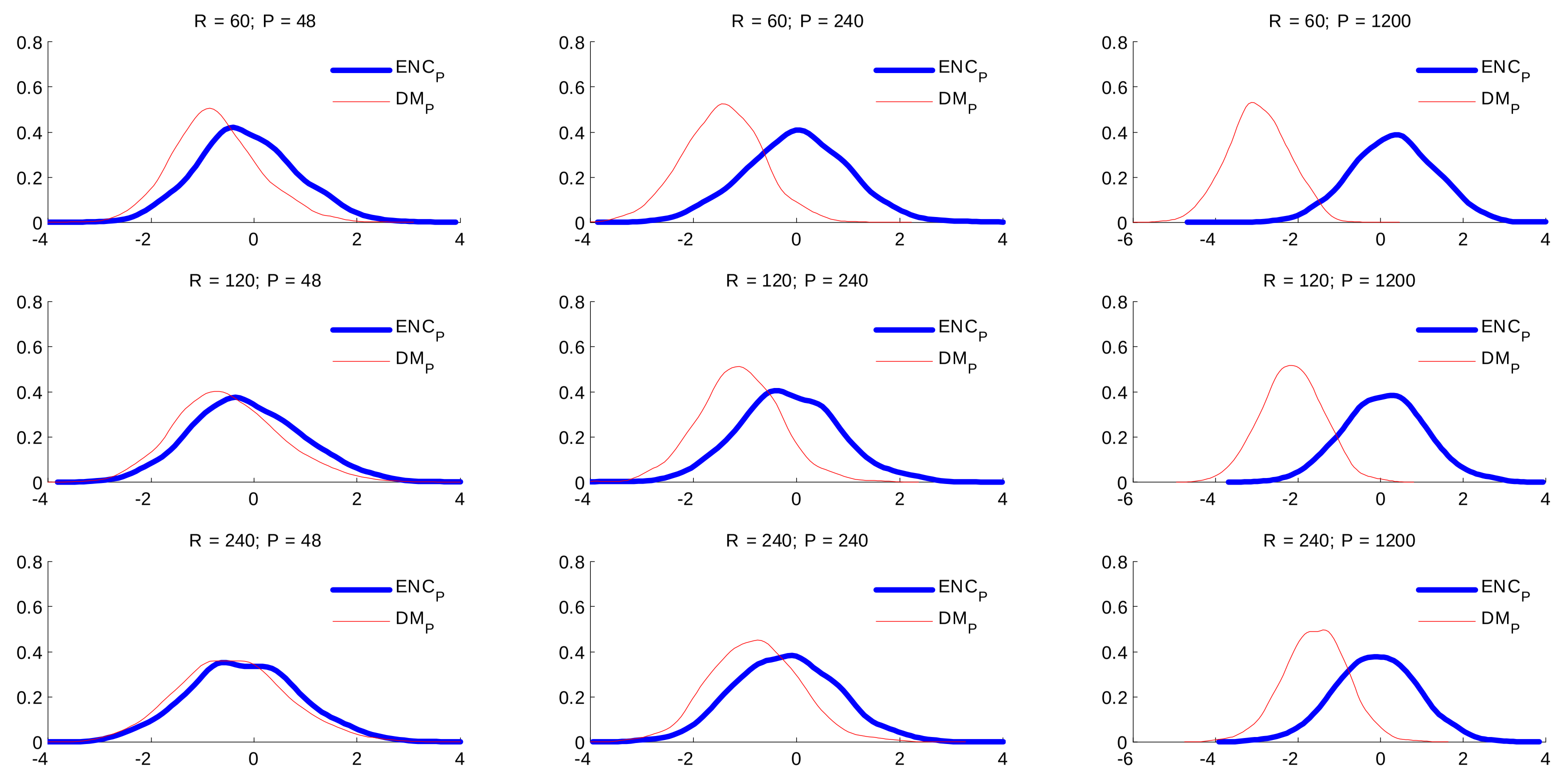

Figure 27: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0.99$, $b = 0$, $\sigma_e = 1$, $\rho(e_t, v_t) = 0.5$, with intercept on Model 1, 2000 Repeats.

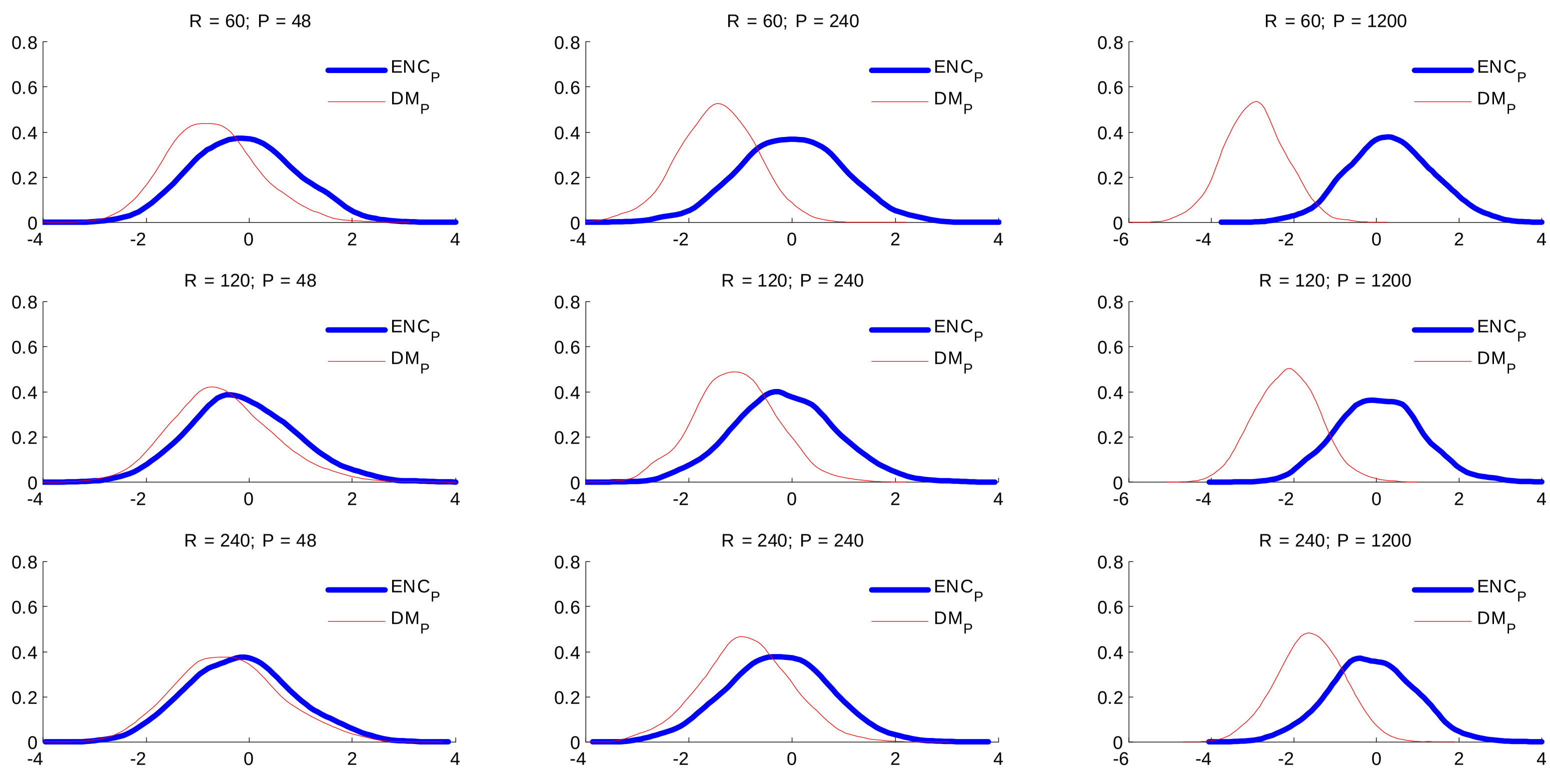

Figure 28: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0.99$, $b = 0$, $\sigma_e = 0.1$, $\rho(e_t, v_t) = 0.5$, with intercept on Model 1, 2000 Repeats.

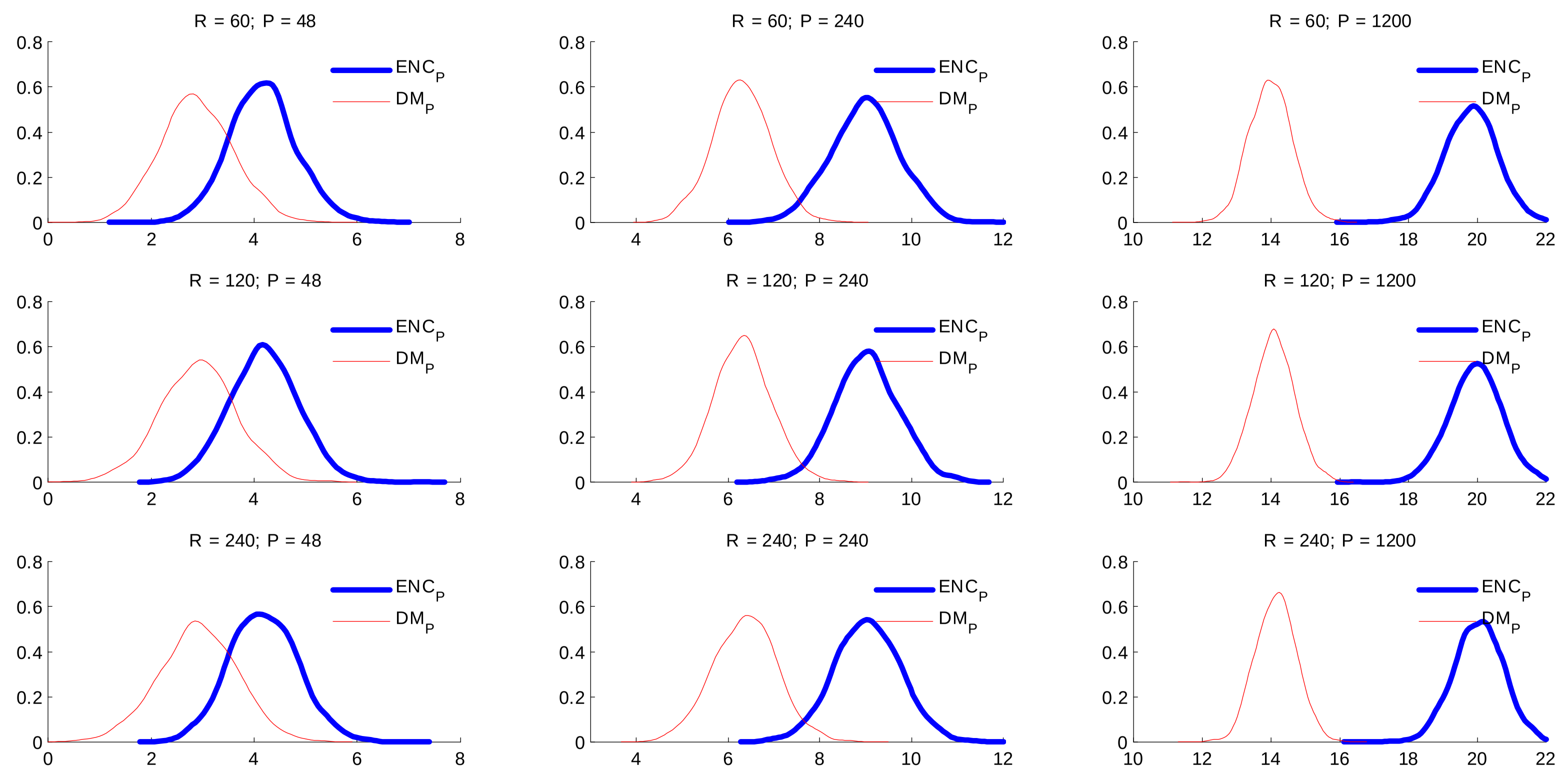

Figure 29: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 0.1$, $\sigma_e = 0.1$, $\rho(e_t, v_t) = 0.5$, with intercept on Model 1, 2000 Repeats.

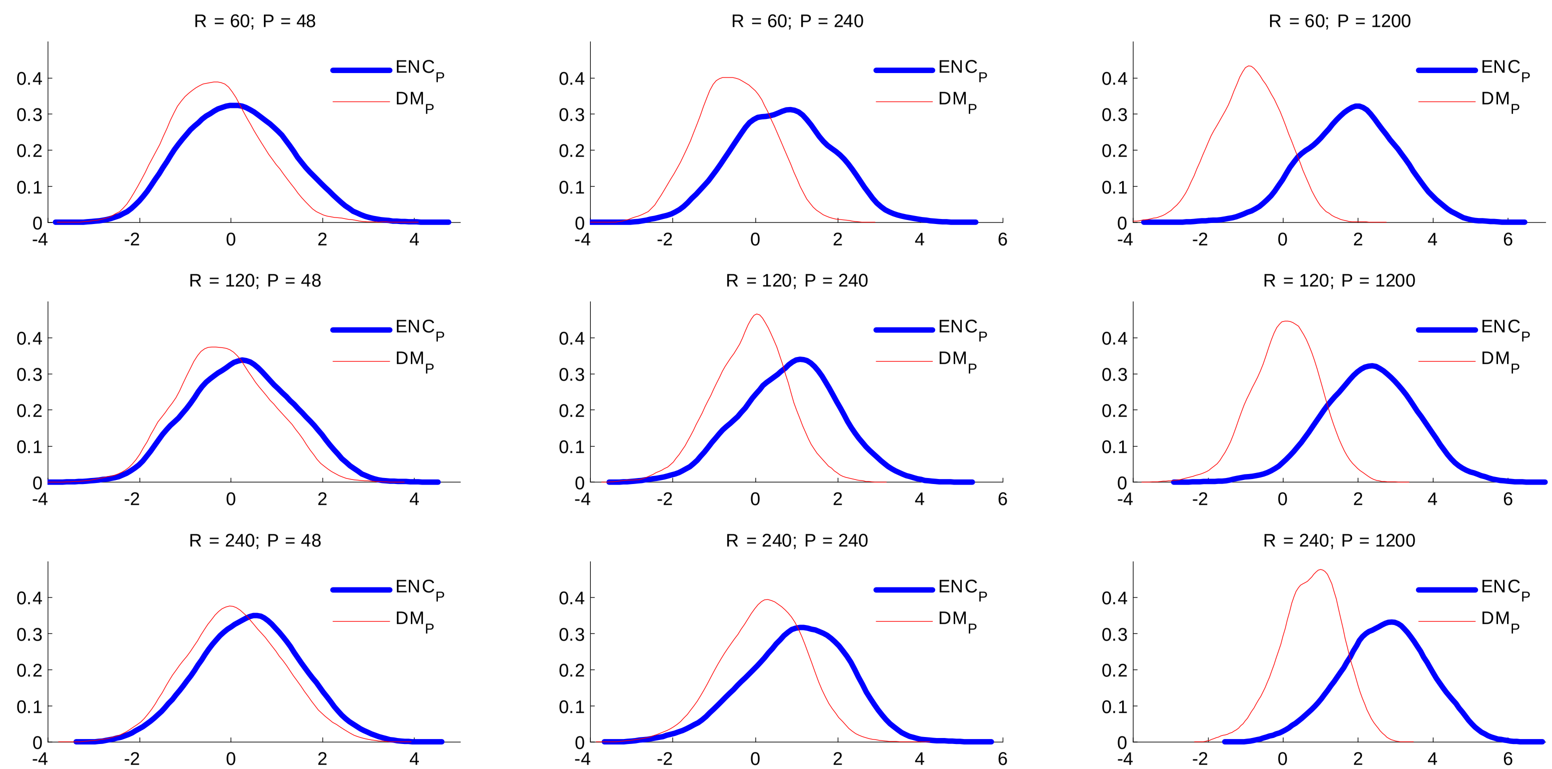

Figure 30: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 0.1$, $\sigma_e = 1$, $\rho(e_t, v_t) = 0.5$, with intercept on Model 1, 2000 Repeats.

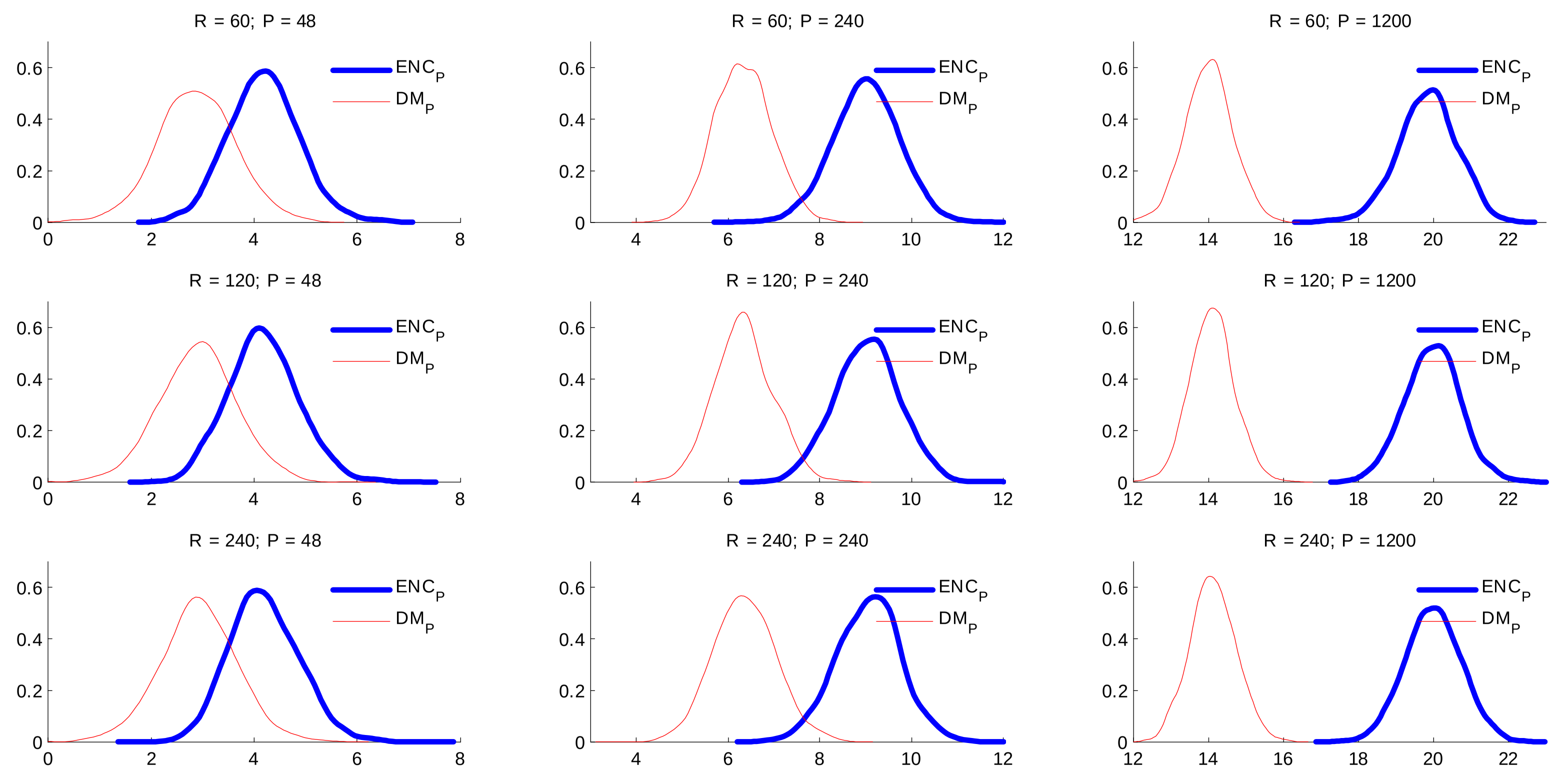

Figure 31: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 1$, $\sigma_e = 1$, $\rho(e_t, v_t) = 0.5$, with intercept on Model 1, 2000 Repeats.

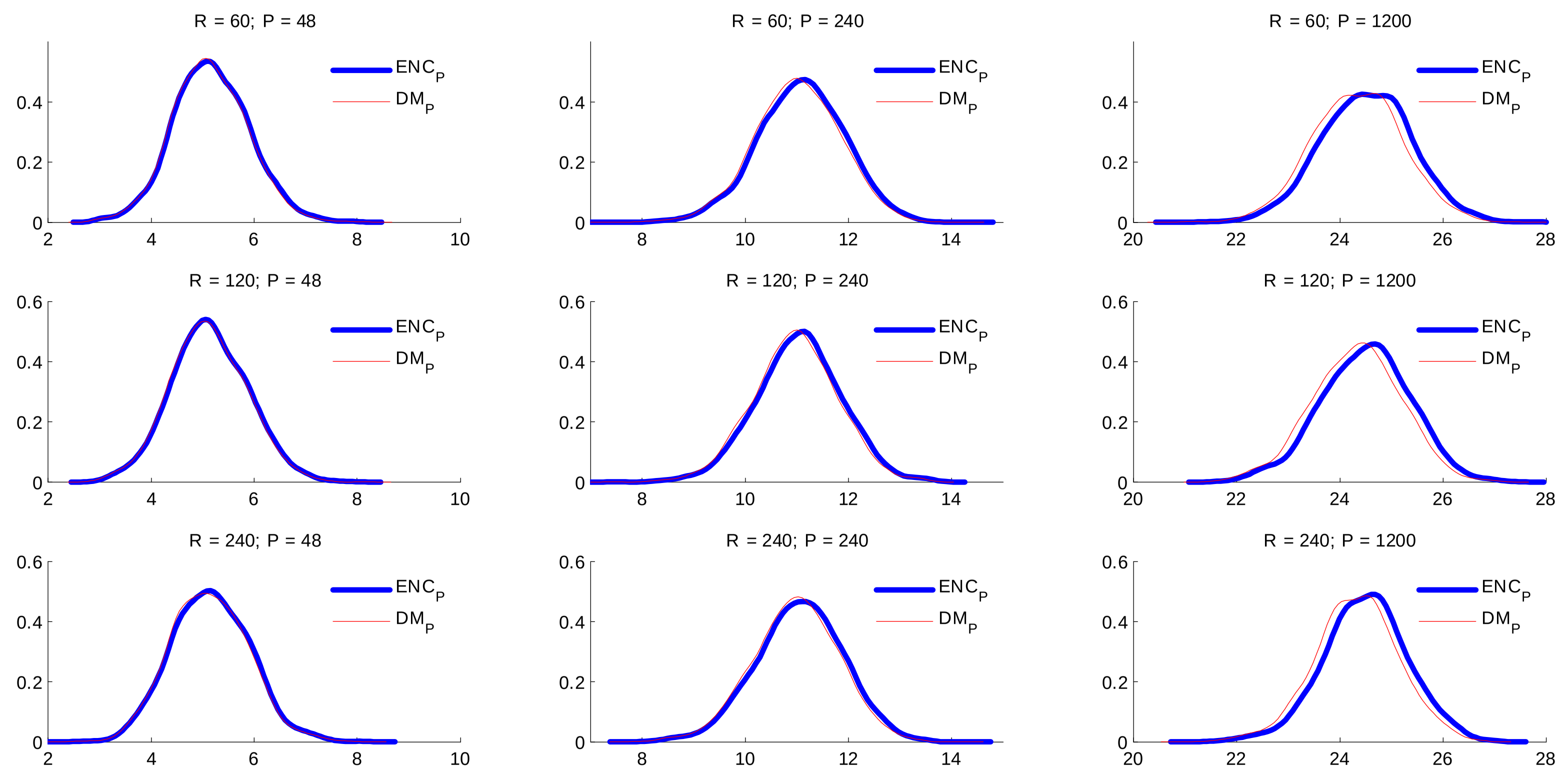

Figure 32: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 1$, $\sigma_e = 0.1$, $\rho\,(e_t, v_t) = 0.5$, with intercept on Model 1, 2000 Repeats.

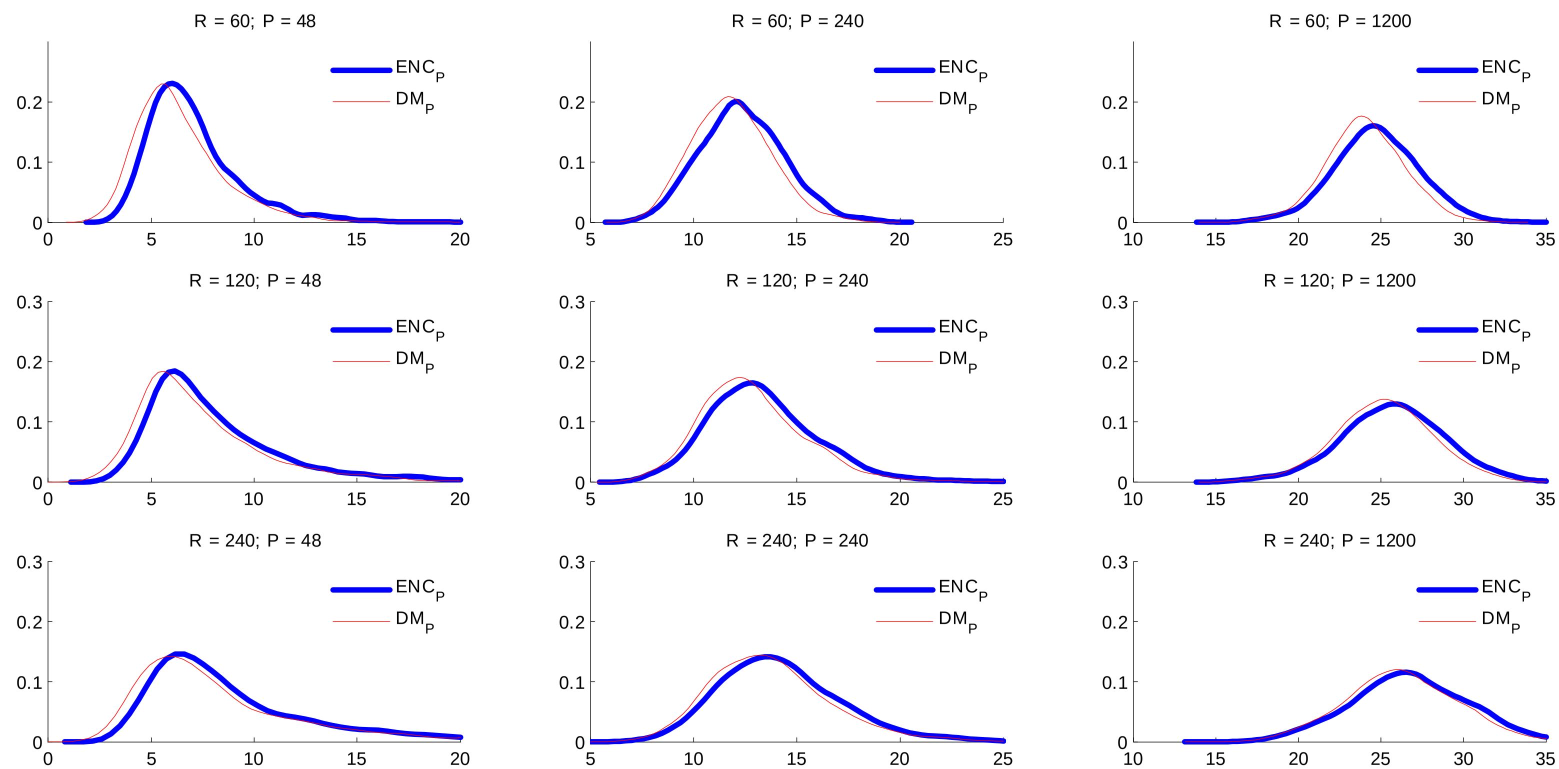

Figure 33: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 0.1$, $\sigma_e = 0.1$, $\rho\left(e_t, v_t\right) = 0.5$, with intercept on Model 1, 2000 Repeats.

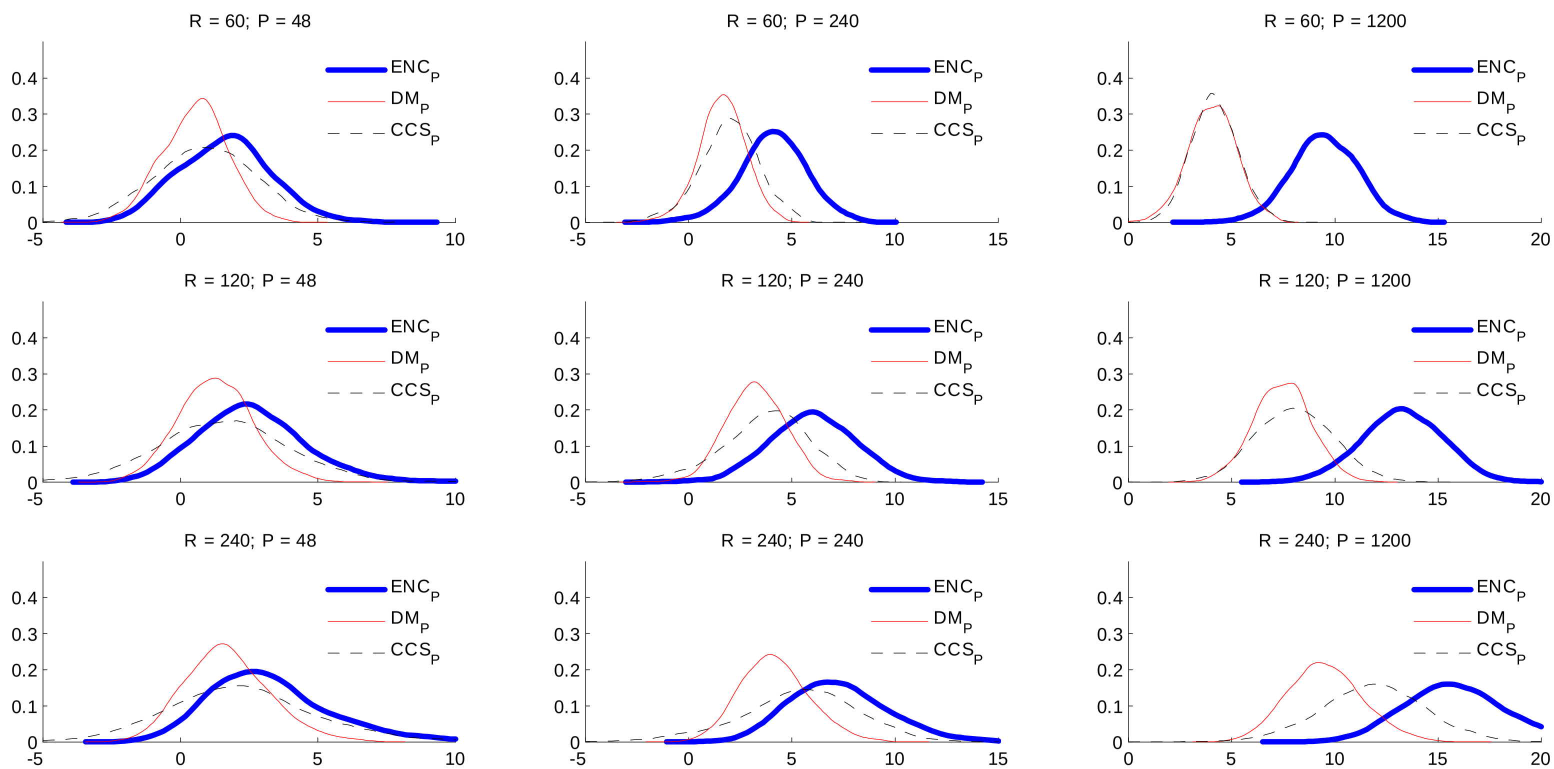

Figure 34: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 0.1$, $\sigma_e = 1$, $\rho\left(e_t, v_t\right) = 0.5$, with intercept on Model 1, 2000 Repeats.

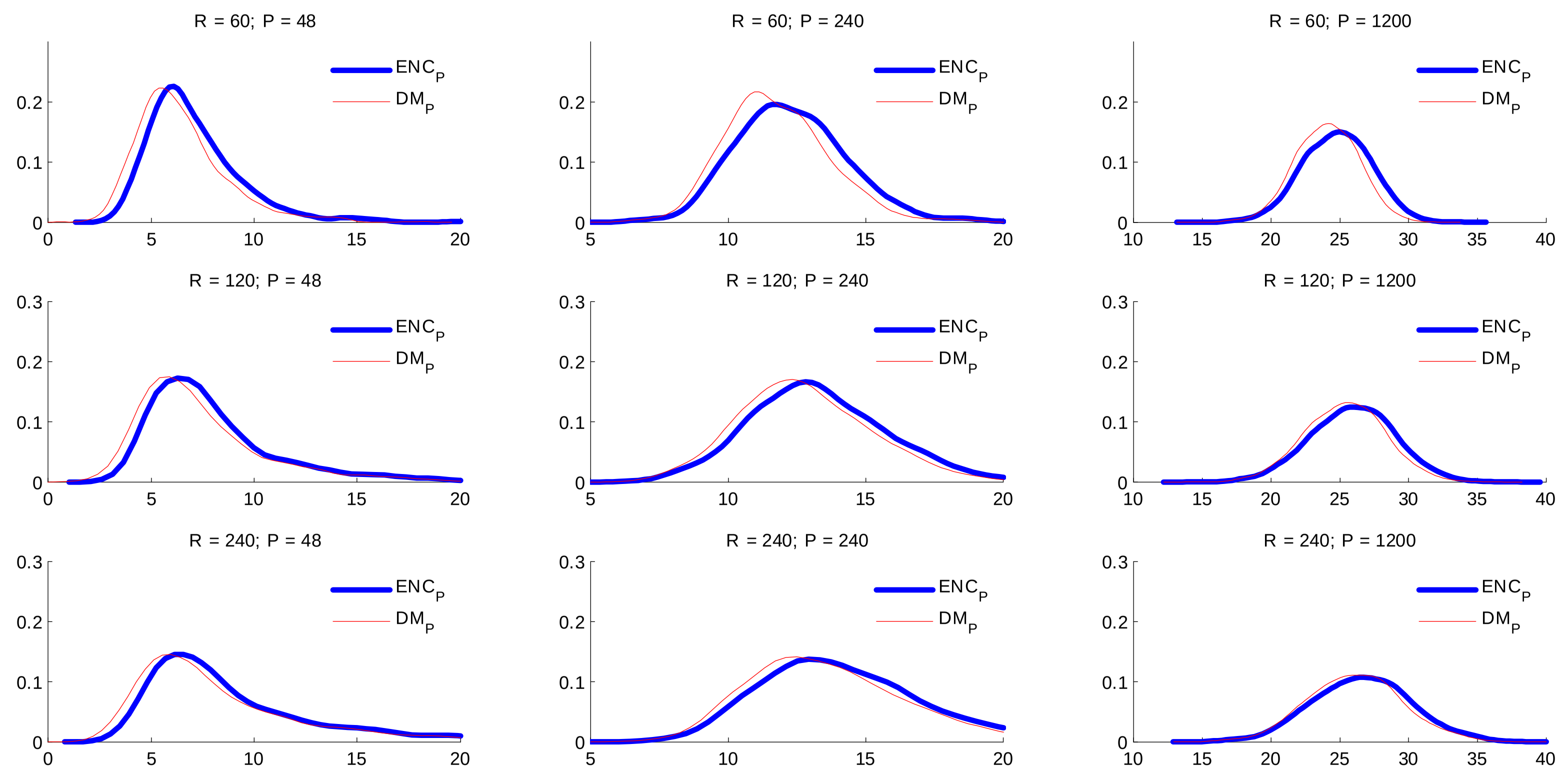

Figure 35: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 1$, $\sigma_e = 1$, $\rho(e_t, v_t) = 0.5$, with intercept on Model 1, 2000 Repeats.

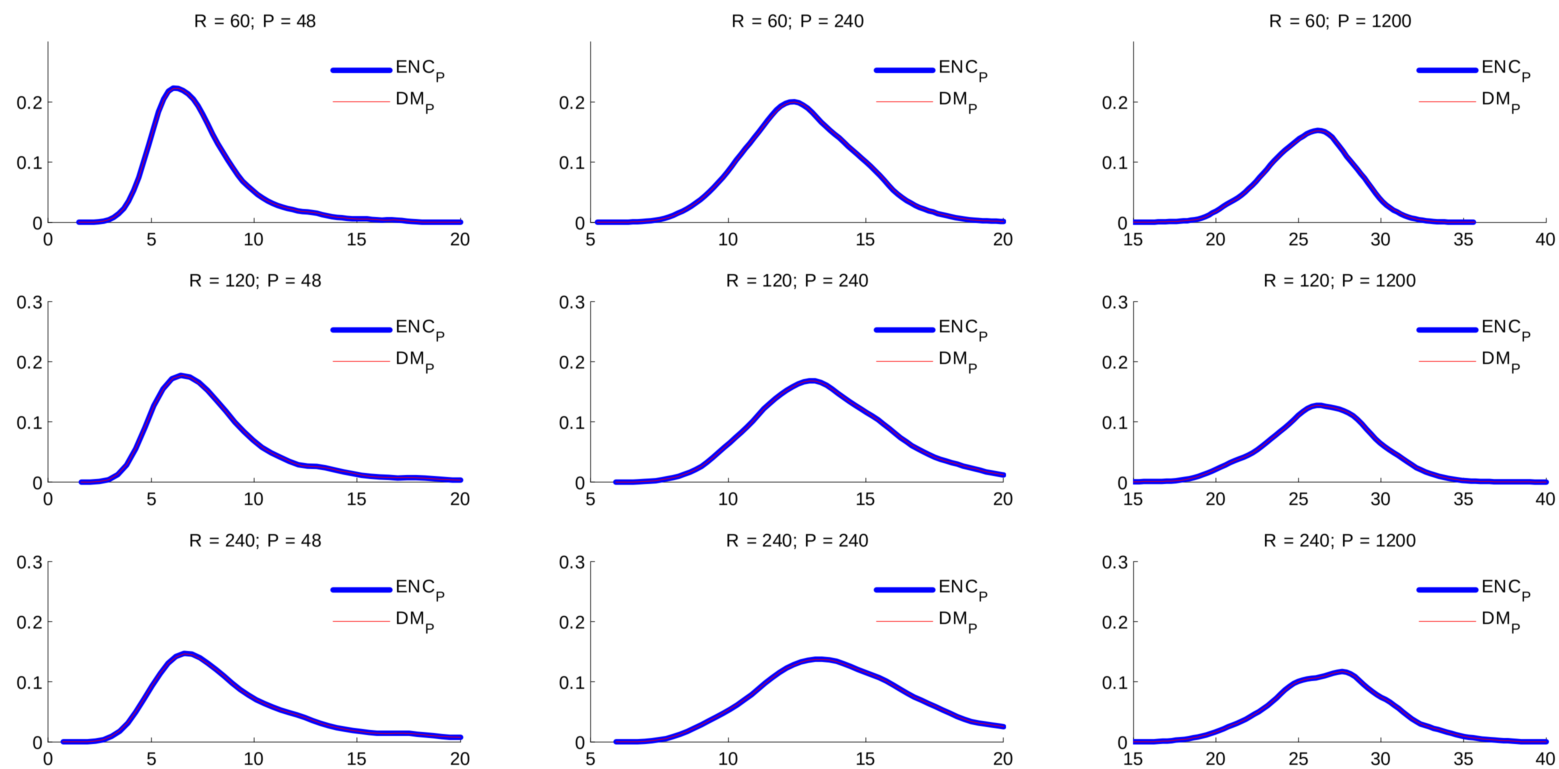

Figure 36: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 1$, $\sigma_e = 0.1$, $\rho(e_t, v_t) = 0.5$, with intercept on Model 1, 2000 Repeats.

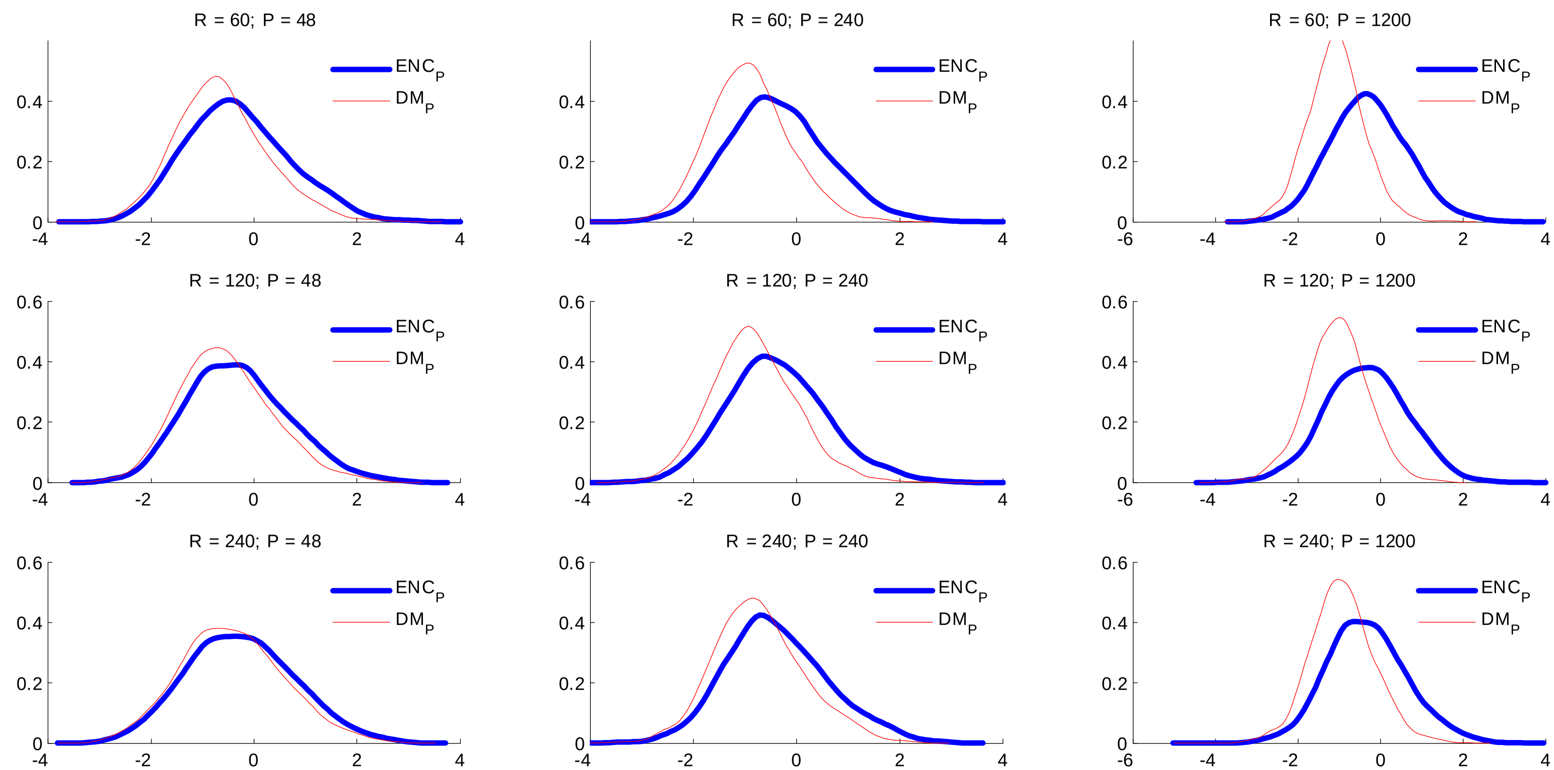

Figure 37: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0$, $b = 0$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

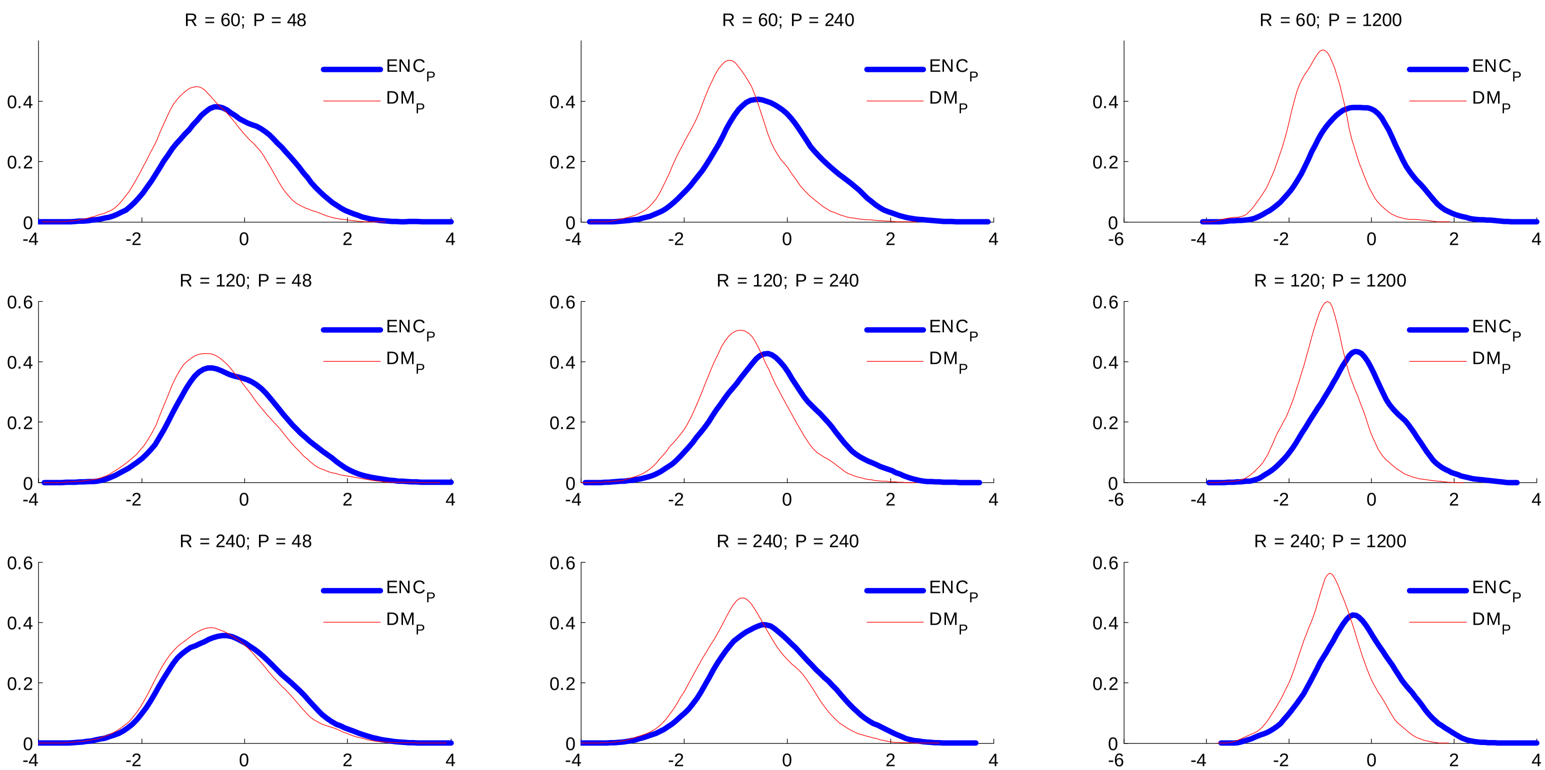

Figure 38: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0$, $b = 0$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

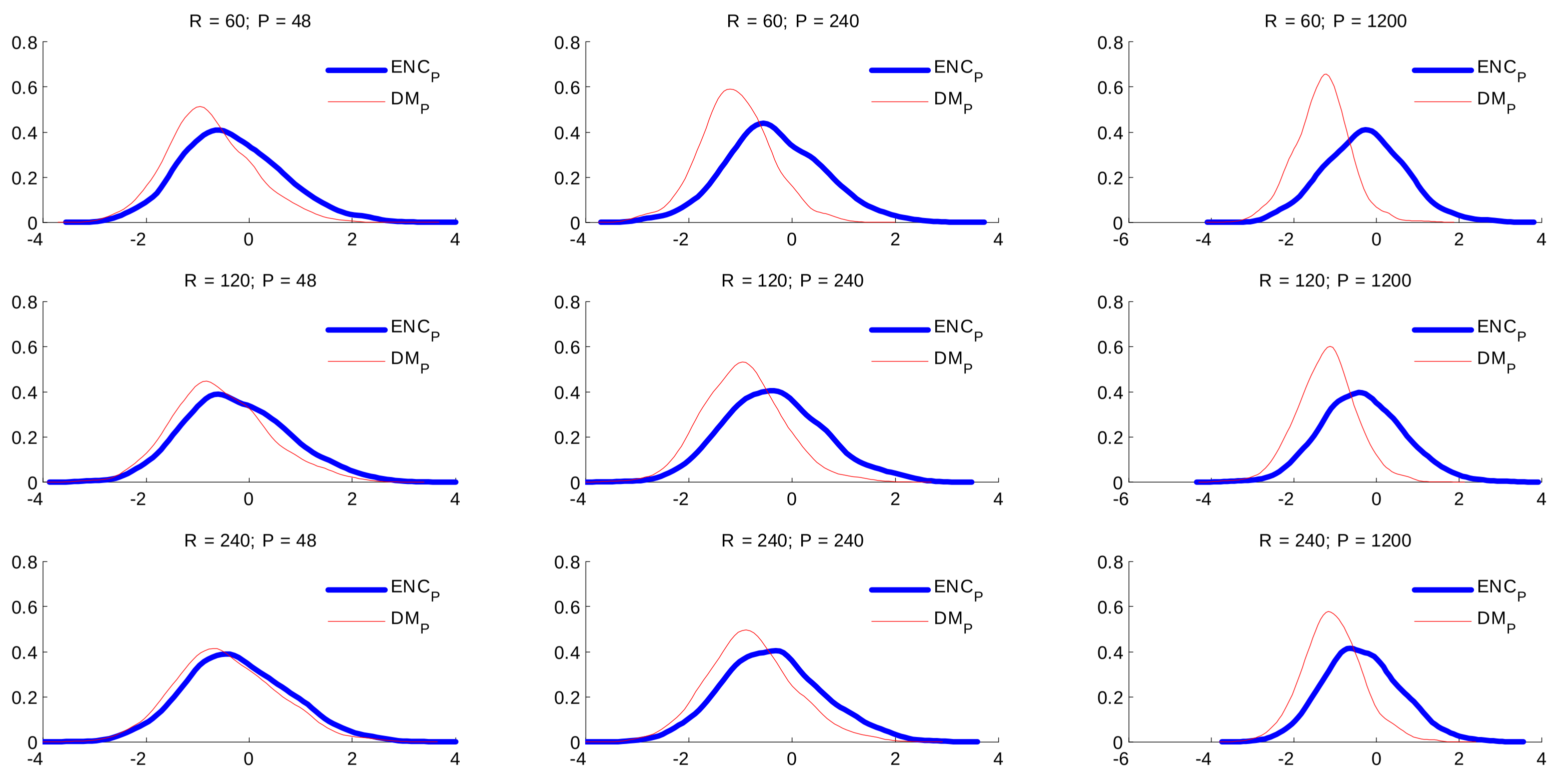

Figure 39: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0.99$, $b = 0$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

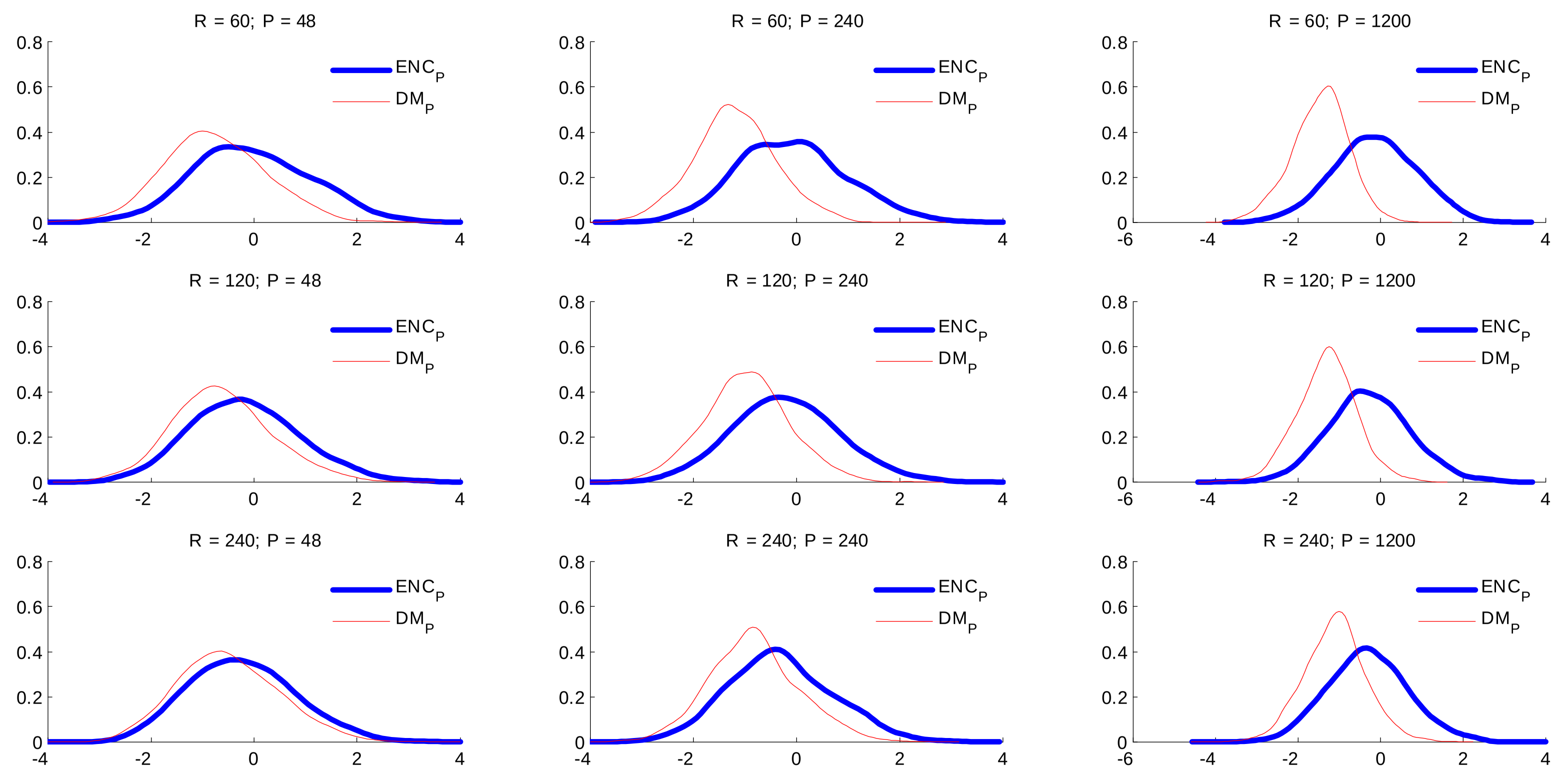

Figure 40: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_0$, $\phi = 0.99$, $b = 0$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

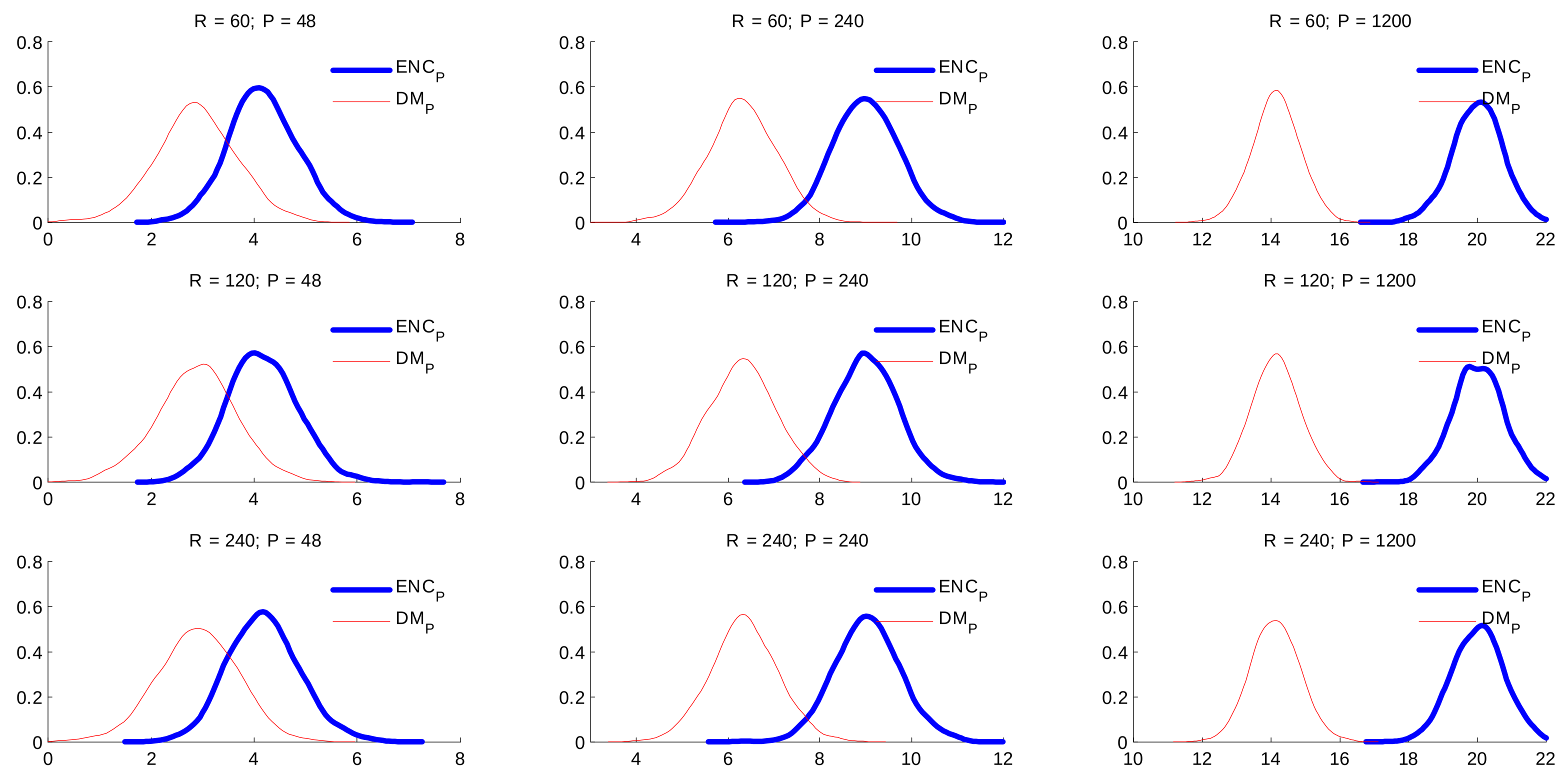

Figure 41: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 0.1$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

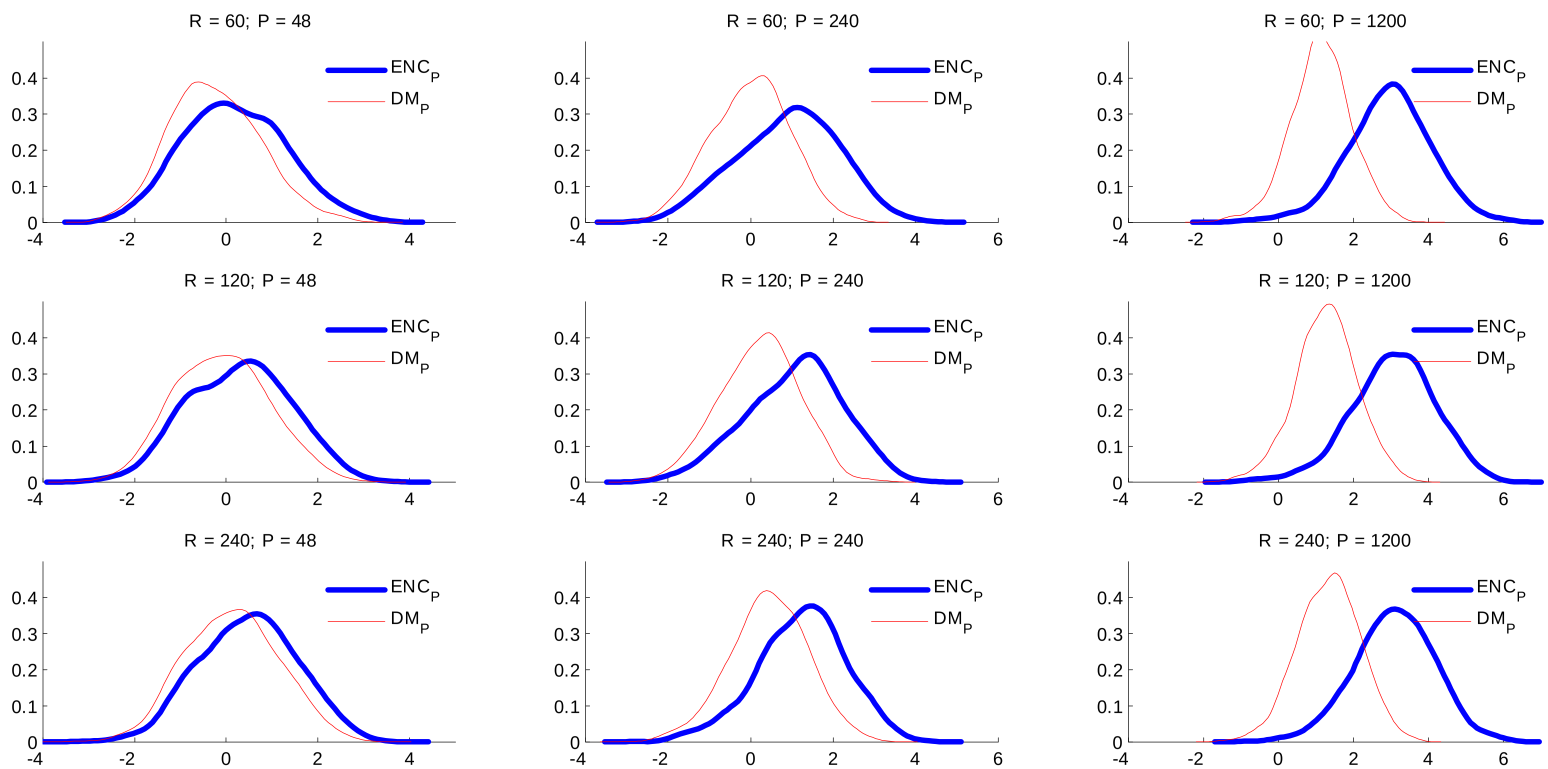

Figure 42: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 0.1$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

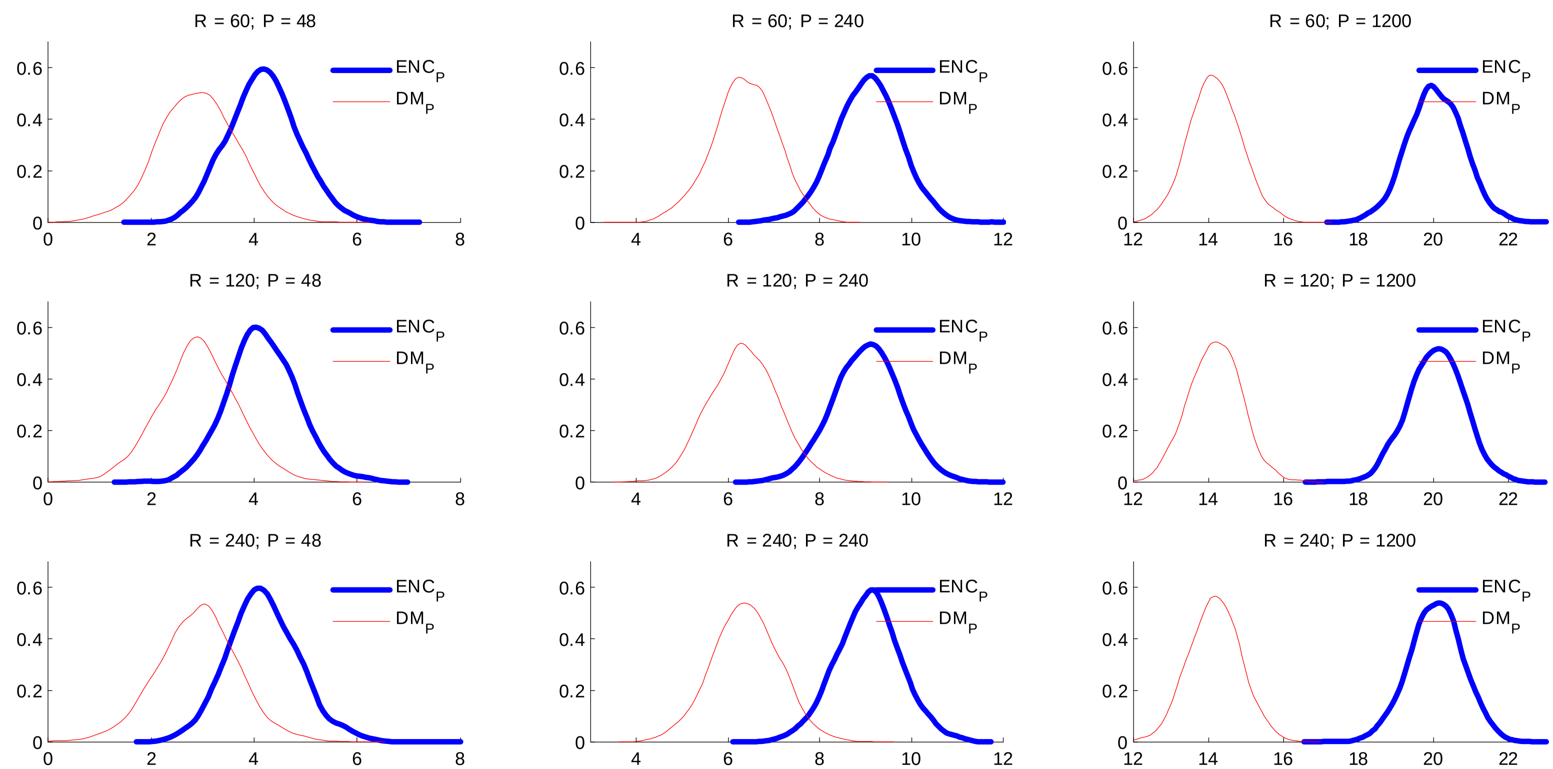

Figure 43: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 1$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

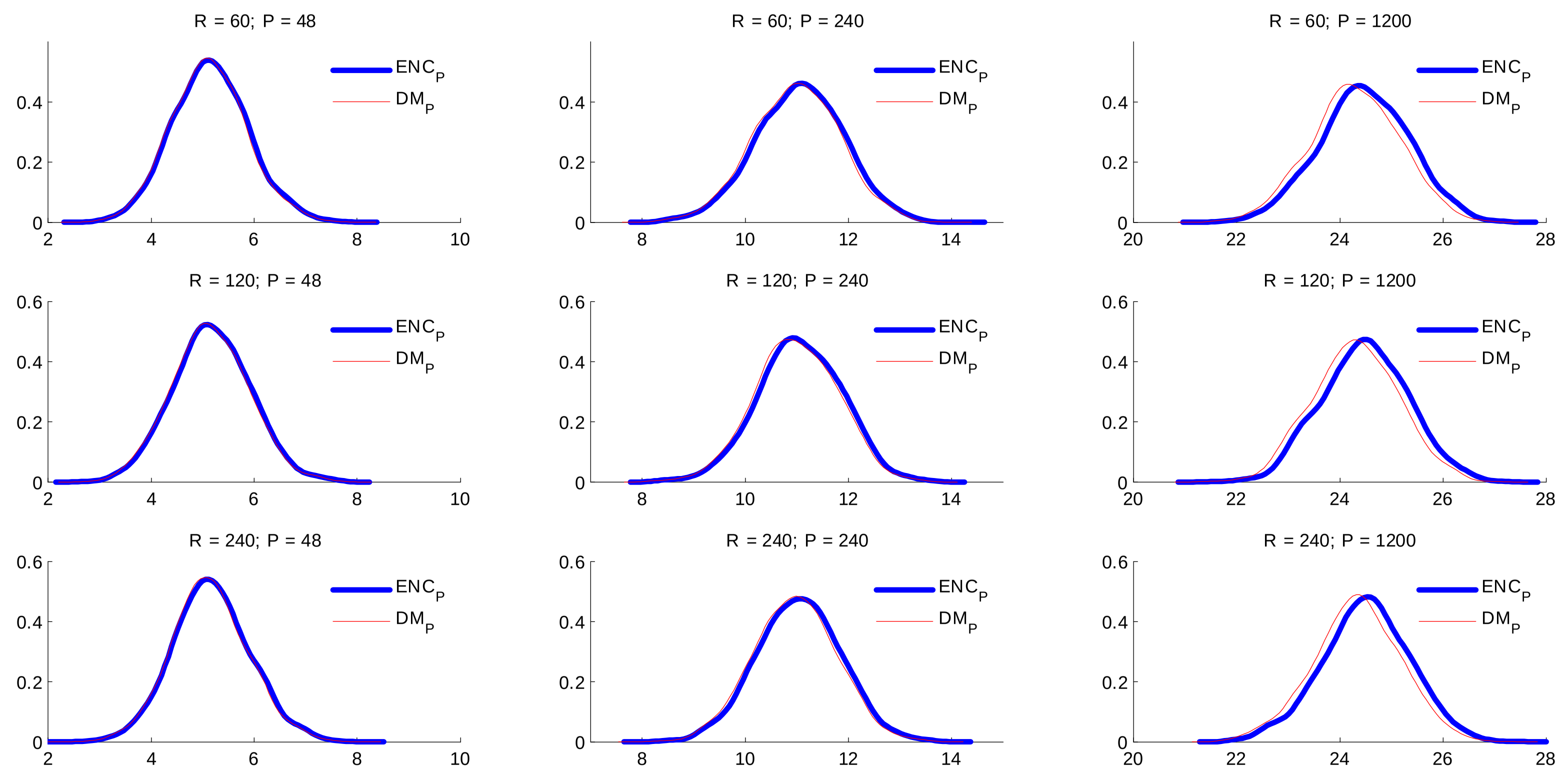

Figure 44: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0$, $b = 1$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

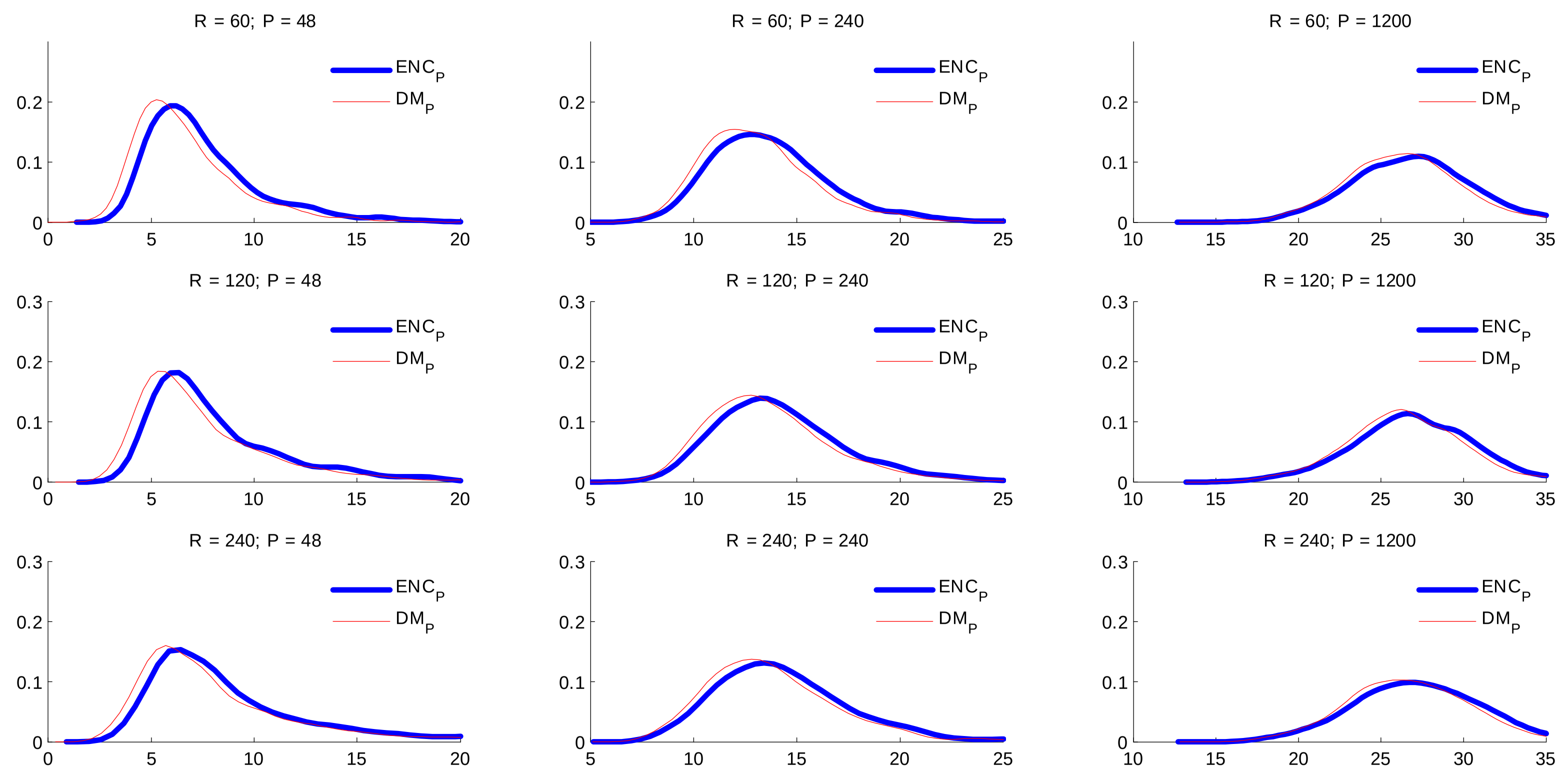

Figure 45: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 0.1$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

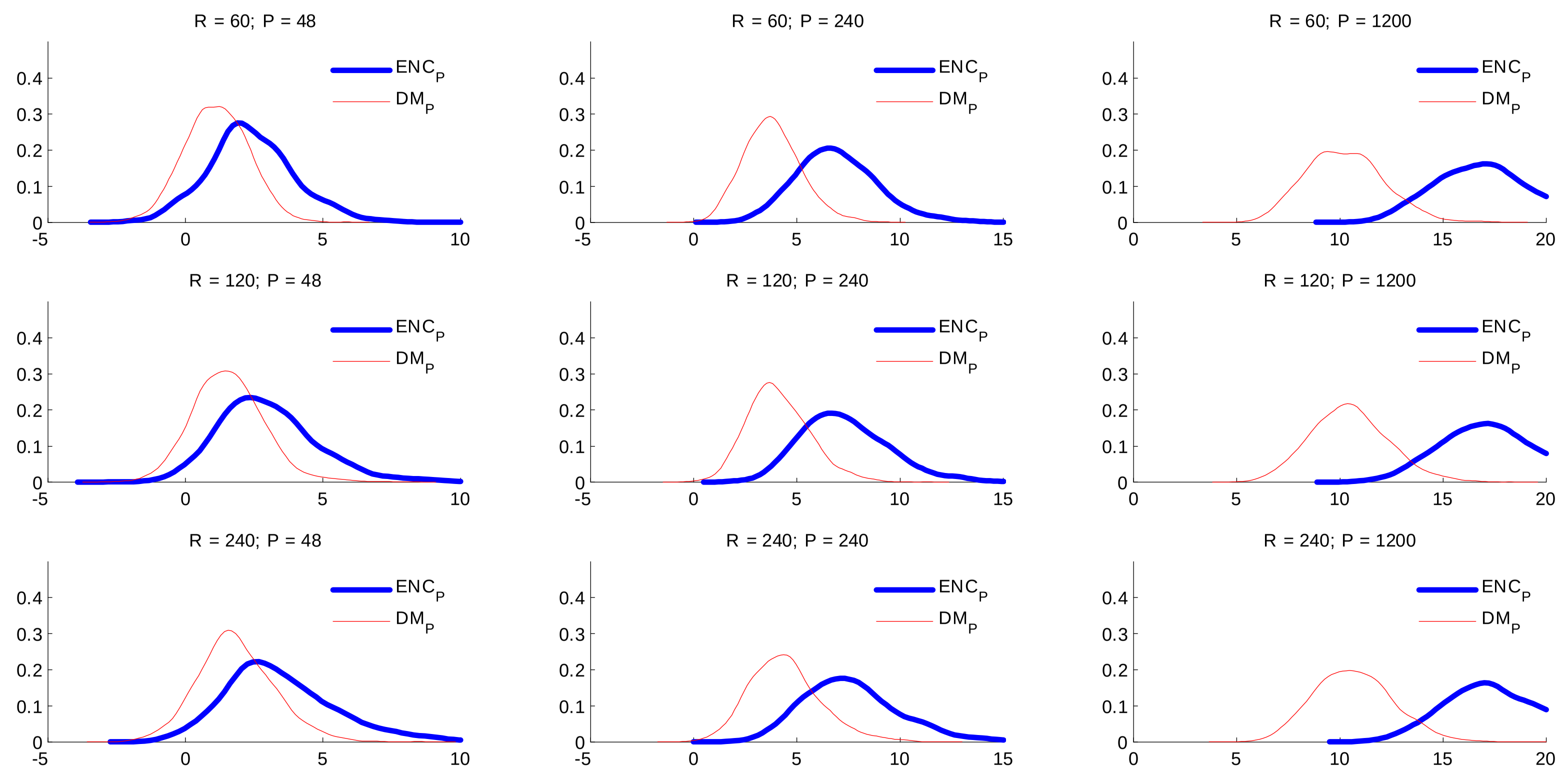

Figure 46: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 0.1$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

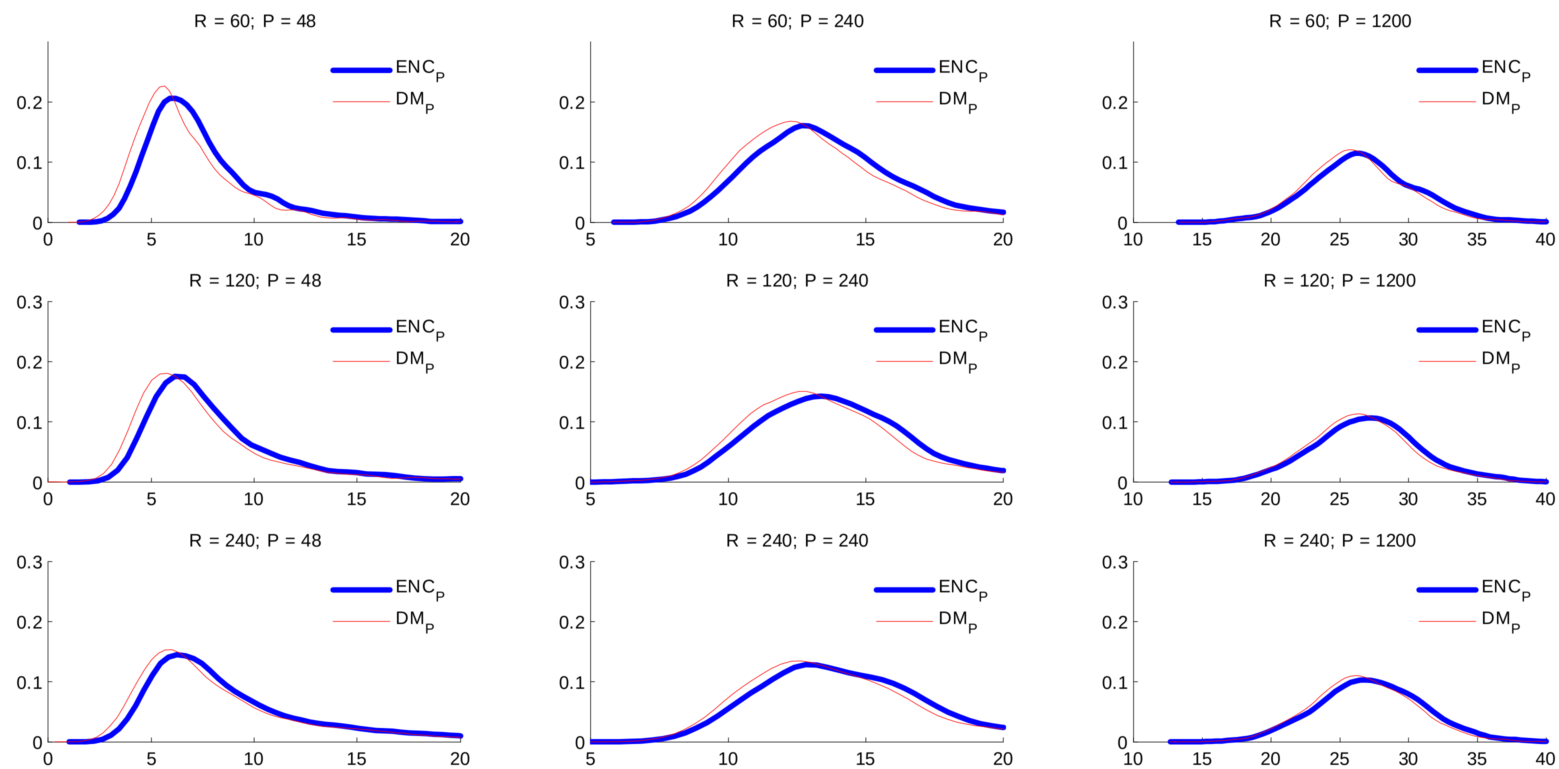

Figure 47: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 1$, $\sigma_e = 1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

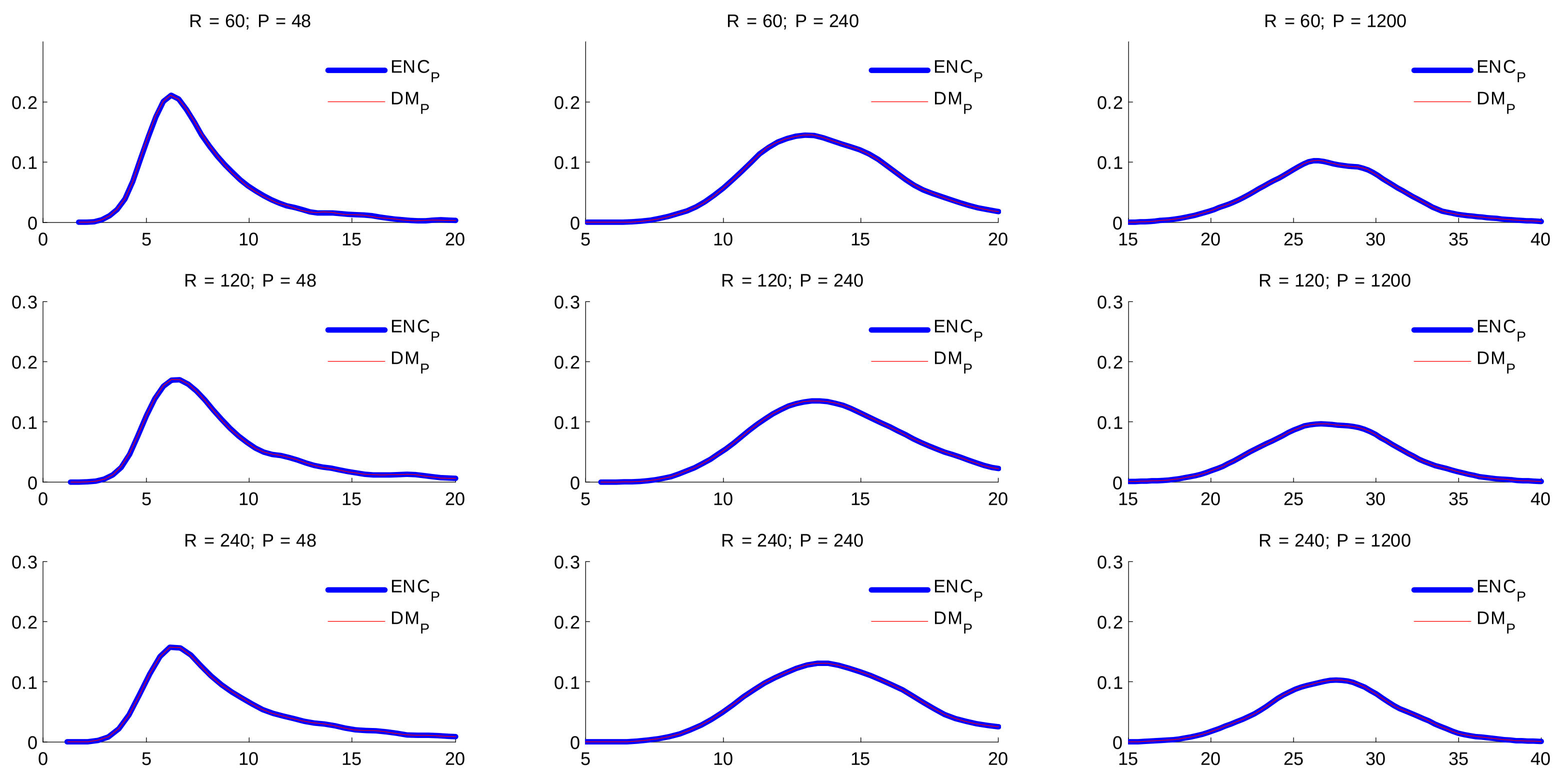

Figure 48: Monte Carlo distribution of ENC (blue line), and DM (red line) under $\mathbb{H}_1$, $\phi = 0.99$, $b = 1$, $\sigma_e = 0.1$, with intercept on Model 1, 2000 Repeats. Recursive Scheme.

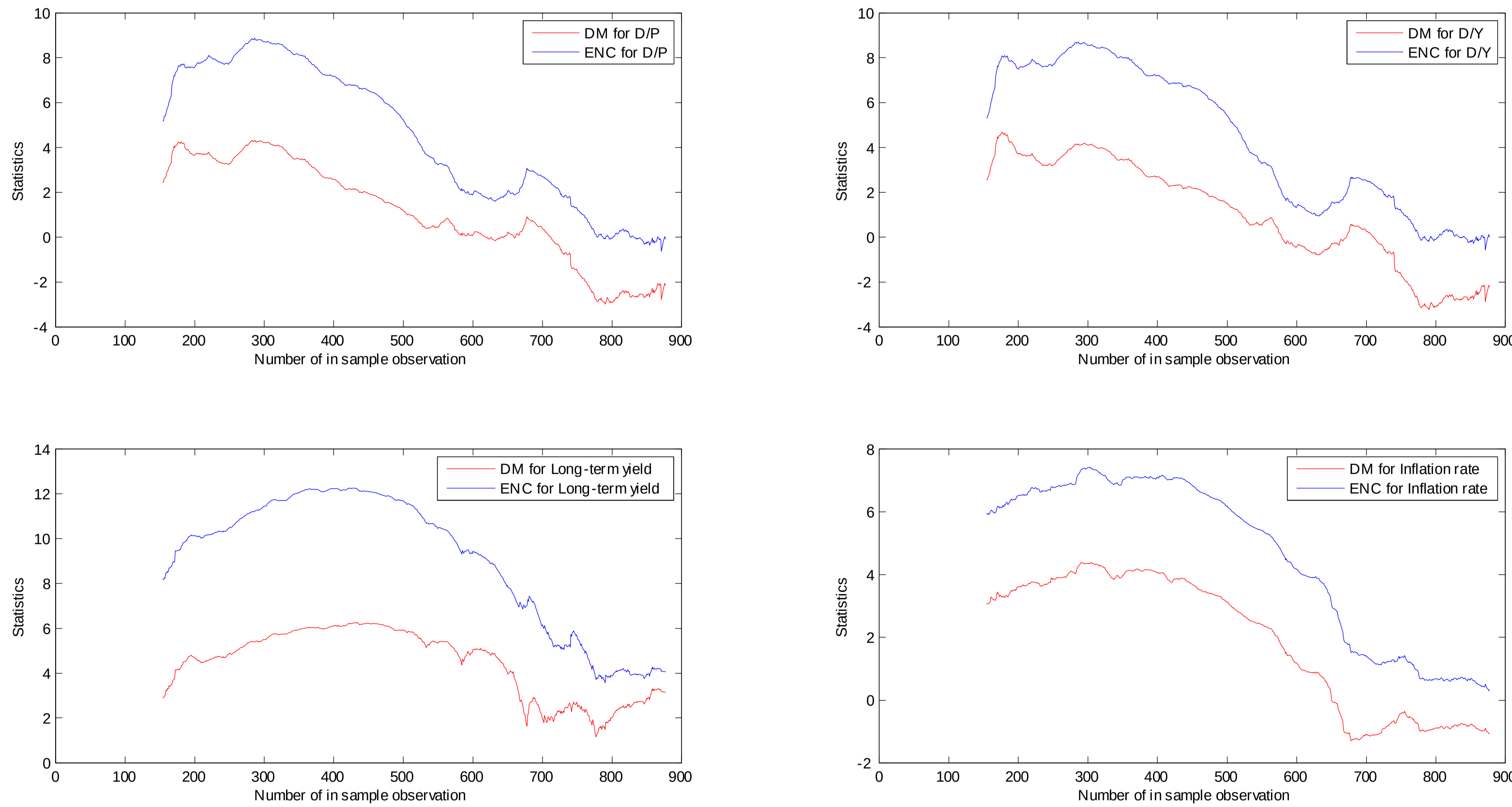

Figure 49: Testing for predictive ability of persistent predictors for equity premium using rolling scheme (small model has constant term)